

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
21 novembre 2002 (21.11.2002)

PCT

(10) Numéro de publication internationale
WO 02/092818 A2

(51) Classification internationale des brevets⁷ :

C12N 15/31, 15/62, 15/74, 5/10,
1/20, 1/21, C07K 14/315, 16/12, 19/00, C12Q 1/68, G01N
33/53, A61K 48/00, 39/395, 39/09, A61P 31/04

(21) Numéro de la demande internationale :

PCT/IB02/03059

(22) Date de dépôt international : 26 avril 2002 (26.04.2002)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :

2001/05642 26 avril 2001 (26.04.2001) FR

(71) Déposants (*pour tous les États désignés sauf US*) : **INSTITUT PASTEUR** [FR/FR]; 25-28, rue du Docteur Roux, F-75015 Paris (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel-Ange, F-75016 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : **GLASER, Philippe** [FR/FR]; 72 rue de la Glacière, F-75013 Paris (FR). **RUSNIOK, Christophe** [FR/FR]; 44 rue des Gallardons, F-92350 Le Plessis Robinson (FR). **CHEVALIER, Fabien** [FR/FR]; 29, avenue Léon Blum, F-94230 Cachan (FR). **FRANGEUL, Lionel** [FR/FR]; 2, rue Mouton Duvernet, F-75014 Paris (FR). **LALIOUI, Lila** [DZ/FR]; 7, rue des Presles, F-94170 Le Perreux sur Marne (FR). **ZOUINE, Mohamed** [FR/FR]; 4, avenue de Normandie, log. 156 91940, F- Les Ulis (FR). **COUVE, Elisabeth** [FR/FR]; 6, rue de Limours, F-78120 Rambouillet (FR). **BUCHRIESER, Carmen** [AT/FR]; 11, rue de l'Amiral

Mouchez, F-75013 Paris (FR). **POYART, Claire** [FR/FR]; 20, rue Léon Blum, F-92260 Fontenay-aux-Roses (FR). **TRIEU-CUOT, Patrick** [FR/FR]; 20, rue Léon Blum, F-92260 Fontenay aux Roses (FR). **KUNST, Frank** [FR/FR]; 46, rue Barbes, F-92200 Ivry sur Seine (FR).

(74) Mandataires : **MARTIN, Jean-Jacques** etc.; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) États désignés (*national*) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée :

- sans rapport de recherche internationale, sera republiée dès réception de ce rapport
- avec la partie réservée au listage des séquences de la description publiée séparément sous forme électronique et disponible sur demande auprès du Bureau international

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) Title: STREPTOCOCCUS AGALACTIAE GENOME SEQUENCE, USE FOR DEVELOPING VACCINES, DIAGNOSTIC TOOLS, AND FOR IDENTIFYING THERAPEUTIC TARGETS

(54) Titre : SEQUENCE DU GENOME STREPTOCOCCUS AGALACTIAE, APPLICATION AU DEVELOPPEMENT DE VACCINS, D'OUTILS DE DIAGNOSTIC, ET A L'IDENTIFICATION DE CIBLES THERAPEUTIQUES

(57) Abstract: The invention concerns the genome sequence and nucleotide sequences coding for *Streptococcus agalactiae* polypeptides, such as cellular envelope polypeptides, or secreted or specific polypeptides, or polypeptides involved in the metabolism and the replication process, as well as vectors or cells comprising said sequences. The invention also concerns the use thereof for developing vaccines, diagnostic tools, DNA chips and for identifying therapeutic targets.

(57) Abrégé : L'invention a pour objet la séquence génomique et des séquences nucléotidiques codant pour des polypeptides de *Streptococcus agalactiae*, tels que des polypeptides d'enveloppe cellulaire, ou des polypeptides sécrétés ou spécifiques, ou impliqués dans le métabolisme et dans le processus de réplication, ainsi que des vecteurs ou cellules incluant lesdites séquences. L'invention porte également sur leur application au développement de vaccins, d'outils de diagnostic, de puces à ADN et à l'identification de cibles thérapeutiques.



WO 02/092818 A2

Séquence du génome *Streptococcus agalactiae*, application au développement de vaccins, d'outils de diagnostic, et à l'identification de cibles thérapeutiques.

5 L'invention a pour objet la séquence génomique et des séquences nucléotidiques codant pour des polypeptides de *Streptococcus agalactiae*, tels que des polypeptides d'enveloppe cellulaire, ou des polypeptides sécrétés ou spécifiques, ou impliqués dans le métabolisme et dans le processus de répllication, ainsi que des vecteurs ou cellules incluant lesdites séquences. L'invention porte également sur leur application au développement de vaccins, d'outils de diagnostic, de puces à ADN et à l'identification
10 de cibles thérapeutiques.

Streptococcus agalactiae est un streptocoque β -hémolytique qui constitue la seule espèce appartenant au groupe B de Lancefield (SGB). Les infections néonatales à streptocoques du groupe B posent un important problème de santé publique qui ne se
15 limite pas aux pays en voie de développement. Leur incidence est de 2,5 pour 1 000 naissances, avec un taux de mortalité qui varie actuellement dans les pays industrialisés entre 4 et 10 % selon les études. Cette bactérie est responsable d'environ 20 % des méningites bactériennes recensées en France et des séquelles neurologiques sont alors observées dans 25 à 50 % des cas. Elle est également à l'origine de mort foetale *in utero*.
20 Le polyoside capsulaire est l'antigène de surface majeur des SGB. Cinq sérotypes (Ia, Ib, II, III et V) sont généralement détectés au cours des infections humaines, le sérotype III étant retrouvé dans 75 % des infections néonatales avec atteinte méningée. Au niveau cellulaire et moléculaire, les différentes étapes du processus infectieux dû à *S. agalactiae* sont encore peu connues. Il est vraisemblable que, dans le cas du syndrome
25 précoce (infections survenant dans les 24 premières heures), la bactérie inhalée pénètre dans les cellules de l'épithélium alvéolaire du nouveau-né et traverse cette barrière pour disséminer ultérieurement dans la circulation générale. La genèse du syndrome tardif précoce (infections survenant entre le 7ème jour et le 3ème mois) et des autres infections à SGB reste encore très mal comprise. Le seul facteur de virulence des SGB
30 dont le rôle a été clairement démontré est le polyoside capsulaire qui permet l'échappement au système immunitaire de l'hôte. La contribution exacte de certaines protéines de surface (antigène C, protéine Rib et C5a peptidase) à la virulence de cette bactérie est encore peu connue.

Une recherche réalisée sur le site EXPASY (<http://www.expasy.ch/>) indique qu'il existe 112 références de séquences protéiques dans les banques Swissprot et TREMBL. Ce nombre inclut des protéines codées par des plasmides de *S. agalactiae*. Ces séquences représentent donc une vision partielle d'un nombre limité d'aspect de la biologie de *S. agalactiae*. La biosynthèse de la capsule polysaccharidique est un des aspects les mieux connus de la virulence de cette bactérie. Par ailleurs, les gènes codant pour 6 protéines exposées à la surface sont également connus (3).

Afin d'appréhender de manière globale les déterminants génétiques impliqués dans ces processus ainsi que le métabolisme de *Streptococcus agalactiae*, le séquençage du génome de *Streptococcus agalactiae* a été réalisé. Le génome de la souche *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) qui a été responsable d'une septicémie mortelle a été choisi pour ce séquençage. Cette souche possède un sérotype capsulaire III, ne présente pas de résistance acquise aux antibiotiques, est génétiquement modifiable et est virulente dans un modèle d'infection murin. La connaissance complète du génome est une étape cruciale pour la caractérisation des gènes impliqués dans le développement du processus infectieux : adhésion et franchissement des structures épithéliales, échappement au système immunitaire et adaptation à des conditions de culture variées et souvent hostiles (pH, stress oxydatif et carences nutritionnelles), qui constituent des cibles potentielles pour de nouvelles stratégies thérapeutiques. La comparaison du génome de *S. agalactiae* avec ceux d'autres pathogènes à Gram positif (*Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans*, *Staphylococcus aureus*, *Listeria monocytogenes*, ...) doit permettre d'identifier de nouveaux gènes de virulence ainsi que de nouvelles cibles pour construire des souches de virulence atténuées et des vaccins. Les protéines de surfaces constituent des candidats pour une future préparation vaccinale. Dans les tableaux 2 et 6 ci-après sont répertoriés respectivement 25 et 30 nouveaux gènes, nouvellement identifiés, codant pour des protéines potentiellement liées au peptidoglycane et présentant le motif de liaison LPXTG.

30

La séquence complète du génome de *Streptococcus agalactiae* (CIP 82.45 (ATCC 12403)) a été obtenue. Ce génome est constitué d'un chromosome long d'environ 2,2 Mb identifié ici sous forme de 138 contigs représentés par les séquences SEQ ID No. 1 à SEQ ID No. 136, SEQ ID No. 138 et SEQ ID No. 139, et d'un plasmide

long de 45 kbases présent dans la souche séquencée représenté par la séquence SEQ ID No. 137. La séquence complète du génome est représentée par la séquence SEQ ID No. 2345.

5 Une liste des phases codantes annotées identifiées par l'analyse des séquences de ces contigs est donnée au tableau 1.

Une liste des phases codantes pour des protéines de surface nouvellement identifiées est donnée au tableau 2 comme indiqué précédemment.

Une liste des phases codantes annotées identifiées par l'analyse de la séquence génomique complète SEQ ID No. 2345 est donnée au tableau 3.

10 Une liste des phases codantes pour des protéines de surface identifiées à partir de l'analyse de la séquence génomique complète SEQ ID No. 2345 est donnée au tableau 6 (protéines liées au peptidoglycane), tableau 8 (lipoprotéines). Tableau 9 (autres protéines de surface), Tableau 10(protéines impliquées dans la biosynthèse des composés polysaccharidiques)

15

La présente invention concerne les séquences nucléotidiques et polypeptidiques de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403).

20 Ainsi, c'est un objet de la présente invention que de caractériser la séquence du génome de *Streptococcus agalactiae*, CIP 82.45 (ATCC 12403) contenu dans la banque génomique préparée à partir du génome de cette souche et déposée à la CNCM le 28 décembre 2000 sous le numéro I-2610, ainsi que de tous les gènes et séquences régulatrices non codantes contenus dans ledit génome.

25 La présente invention concerne donc une séquence nucléotidique isolée et/ou purifiée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi les séquences SEQ ID No. 1 à SEQ ID No. 139 et la séquence SEQ ID No. 2345.

La présente invention concerne également une séquence nucléotidique isolée et/ou purifiée, issue de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi :

30 a) une séquence nucléotidique comportant au moins 75 %, 80 %, 85 %, 90 %, 95 % ou 98 % d'identité avec une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345 ;

b) une séquence nucléotidique hybridant dans des conditions de forte stringence avec une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345,

- et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ;
- c) une séquence nucléotidique complémentaire d'une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, ou complémentaire d'une séquence nucléotidique telle que définie en a), ou b), ou une séquence nucléotidique de l'ARN correspondant à l'une des séquences a) ou b) ;
- d) une séquence nucléotidique d'un fragment représentatif d'une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, ou d'un fragment représentatif d'une séquence nucléotidique telle que définie en a), b) ou c) et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ;
- e) une séquence nucléotidique comprenant une séquence telle que définie en a), b), c) ou d) ; et
- f) une séquence nucléotidique telle que définie en a), b), c), d) ou e) modifiée, de préférence comportant au plus 10 %, 5 %, 1 % ou 0, 5 % de nucléotides modifiés par rapport à la séquence de référence.

De façon plus particulière, la présente invention a également pour objet les séquences nucléotidiques isolées et/ou purifiées, caractérisées en ce qu'elles sont issues de SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, et en ce qu'elles codent pour un polypeptide choisi parmi les polypeptides de séquence SEQ ID No. 140 à SEQ ID No. 2344, et SEQ ID No. 2346 à SEQ ID No. 4481.

La présente invention concerne aussi de façon plus générale les séquences nucléotidiques issues de SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, et codant pour un polypeptide de *Streptococcus agalactia*, telles qu'elles peuvent être isolées à partir de SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345.

De plus, les séquences nucléotidiques isolées et/ caractérisées en ce qu'elles comprennent une séquence nucléotidique choisie parmi :

- a) une séquence nucléotidique codant pour un polypeptide choisi parmi les séquences SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481 ;
- b) une séquence nucléotidique comportant au moins 75 %, 80 %, 85 %, 90 %, 95 % ou 98 % d'identité avec une séquence nucléotidique codant pour un polypeptide choisi parmi les séquences SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481 ;

- c) une séquence nucléotidique s'hybridant dans des conditions de forte stringence avec une séquence nucléotidique codant pour un polypeptide, choisi parmi les séquences SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481, et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ;
- d) une séquence nucléotidique complémentaire ou d'ARN correspondant à une séquence telle que définie en a), b) ou c) ;
- e) une séquence nucléotidique d'un fragment représentatif d'une séquence telle que définie en a), b), c) ou d) et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ; et
- f) une séquence telle que définie en a), b), c), d) ou e) modifiée, de préférence comportant au plus 10 %, 5 %, 1 % ou 0,5 % de nucléotides modifiés par rapport à la séquence de référence,
- sont également des objets de l'invention.

Selon une réalisation avantageuse, l'invention a pour objet les séquences nucléotidiques isolées et/ caractérisées en ce qu'elles comprennent une séquence nucléotidique choisie parmi :

- a) une séquence nucléotidique SEQ ID No. 4482 à SEQ ID No. 6617 ;
- b) une séquence nucléotidique comportant au moins 75 %, 80 %, 85 %, 90 %, 95 % ou 98 % d'identité avec une séquence nucléotidique choisi parmi les séquences SEQ ID No. 2346 à SEQ ID No. 4481 ;
- c) une séquence nucléotidique s'hybridant dans des conditions de forte stringence avec une séquence nucléotidique choisi parmi les séquences SEQ ID No. 4482 à SEQ ID No. 6617, et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ;
- d) une séquence nucléotidique complémentaire ou d'ARN correspondant à une séquence telle que définie en a), b) ou c) ;
- e) une séquence nucléotidique d'un fragment représentatif d'une séquence telle que définie en a), b), c) ou d) et comprenant au moins 20 nucléotides, de préférence 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000 ou 1500 nucléotides ; et

f) une séquence telle que définie en a), b), c), d) ou e) modifiée, de préférence comportant au plus 10 %, 5 %, 1 % ou 0,5 % de nucléotides modifiés par rapport à la séquence de référence,

5 Par acide nucléique, séquence nucléique ou d'acide nucléique, polynucléotide, oligonucléotide, séquence de polynucléotide, séquence nucléotidique, termes qui seront employés indifféremment dans la présente description, on entend désigner un enchaînement précis de nucléotides, modifiés ou non, permettant de définir un fragment ou une région d'un acide nucléique, comportant ou non des nucléotides non naturels, et
10 pouvant correspondre aussi bien à un ADN double brin, un ADN simple brin qu'à des produits de transcription desdits ADNs. Ainsi, les séquences nucléiques selon l'invention englobent également les PNA (Peptid Nucleic Acid).

Il doit être compris que la présente invention ne concerne pas les séquences nucléotidiques dans leur environnement chromosomique naturel, c'est-à-dire à l'état
15 naturel. Il s'agit de séquences qui ont été isolées et/ou purifiées, c'est-à-dire qu'elles ont été prélevées directement ou indirectement, par exemple par copie, leur environnement ayant été au moins partiellement modifié. On entend ainsi également désigner les acides nucléiques obtenus par synthèse chimique.

Par « pourcentage d'identité » entre deux séquences d'acides nucléiques ou
20 d'acides aminés au sens de la présente invention, on entend désigner un pourcentage de nucléotides ou de résidus d'acides aminés identiques entre les deux séquences à comparer, obtenu après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux séquences étant réparties au hasard et sur toute leur longueur. On entend désigner par "meilleur alignement" ou "alignement
25 optimal", l'alignement pour lequel le pourcentage d'identité déterminé comme ci-après est le plus élevé. Les comparaisons de séquences entre deux séquences d'acides nucléiques ou d'acides aminés sont traditionnellement réalisées en comparant ces séquences après les avoir alignées de manière optimale, ladite comparaison étant réalisée par segment ou par « fenêtre de comparaison » pour identifier et comparer les
30 régions locales de similarité de séquence. L'alignement optimal des séquences pour la comparaison peut être réalisé, outre manuellement, au moyen de l'algorithme d'homologie locale de Smith et Waterman (1981, Ad. App. Math. 2:482), au moyen de l'algorithme d'homologie locale de Needleman et Wunsch (1970, J. Mol. Biol. 48:443), au moyen de la méthode de recherche de similarité de Pearson et Lipman (1988, Proc.

Natl. Acad. Sci. USA 85:2444), au moyen de logiciels informatiques utilisant ces algorithmes (GAP, BESTFIT, BLAST P, BLAST N, FASTA et TFASTA dans le Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI). Afin d'obtenir l'alignement optimal, on utilise de préférence le
5 programme BLAST, avec la matrice BLOSUM 62. On peut également utiliser les matrices PAM ou PAM250.

Le pourcentage d'identité entre deux séquences d'acides nucléiques ou d'acides aminés est déterminé en comparant ces deux séquences alignées de manière optimale, la séquence d'acides nucléiques ou d'acides aminés à comparer pouvant comprendre des
10 additions ou des délétions par rapport à la séquence de référence pour un alignement optimal entre ces deux séquences. Le pourcentage d'identité est calculé en déterminant le nombre de positions identiques pour lesquelles le nucléotide ou le résidu d'acide aminé est identique dans les deux séquences, en divisant ce nombre de positions identiques par le nombre total de positions comparées et en multipliant le résultat
15 obtenu par 100 pour obtenir le pourcentage d'identité entre ces deux séquences.

Par séquences nucléiques présentant un pourcentage d'identité d'au moins 75 %, de préférence 80 %, 85 % ou 90 %, de façon plus préférée 95 % voire 98 %, après alignement optimal avec une séquence de référence, on entend désigner les séquences nucléiques présentant, par rapport à la séquence nucléique de référence, certaines
20 modifications comme en particulier une délétion, une troncation, un allongement, une fusion chimérique et/ou une substitution, notamment ponctuelle, et dont la séquence nucléique présente au moins 75 %, de préférence 80 %, 85 %, 90 %, 95 % ou 98 %, d'identité après alignement optimal avec la séquence nucléique de référence. Il s'agit de préférence de séquences dont les séquences complémentaires sont susceptibles de
25 s'hybrider spécifiquement avec les séquences de référence. De préférence, les conditions d'hybridation spécifiques ou de forte stringence seront telles qu'elles assurent au moins 75 %, de préférence 80 %, 85 %, 90 %, 95 % ou 98 % d'identité après alignement optimal entre l'une des deux séquences et sa séquence complémentaire.

30 Une hybridation dans des conditions de forte stringence signifie que les conditions de température et de force ionique sont choisies de telle manière qu'elles permettent le maintien de l'hybridation entre deux fragments d'ADN complémentaires. A titre illustratif, des conditions de forte stringence de l'étape d'hybridation aux fins de

définir les fragments polynucléotidiques décrits ci-dessus, sont avantageusement les suivantes.

L'hybridation ADN-ADN ou ADN-ARN est réalisée en deux étapes : (1) préhybridation à 42°C pendant 3 heures en tampon phosphate (20 mM, pH 7,5) contenant 5 x SSC (1 x SSC correspond à une solution 0,15 M NaCl + 0,015 M citrate de sodium), 50 % de formamide, 7 % de sodium dodécyl sulfate (SDS), 10 x Denhardt's, 5 % de dextran sulfate et 1 % d'ADN de sperme de saumon ; (2) hybridation proprement dite pendant 20 heures à une température dépendant de la taille de la sonde (i.e. : 42°C, pour une sonde de taille > 100 nucléotides) suivie de 2 lavages de 20 minutes à 20°C en 2 x SSC + 2 % SDS, 1 lavage de 20 minutes à 20°C en 0,1 x SSC + 0,1 % SDS. Le dernier lavage est pratiqué en 0,1 x SSC + 0,1 % SDS pendant 30 minutes à 60°C pour une sonde de taille > 100 nucléotides. Les conditions d'hybridation de forte stringence décrites ci-dessus pour un polynucléotide de taille définie, peuvent être adaptées par l'homme du métier pour des oligonucléotides de taille plus grande ou plus petite, selon l'enseignement de Sambrook et al., (1989, Molecular cloning : a laboratory manual. 2nd Ed. Cold Spring Harbor).

De plus, par fragment représentatif de séquences selon l'invention, on entend désigner tout fragment nucléotidique présentant au moins 15 nucléotides, de préférence au moins 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 400, 450, 500, 750, 1000 ou 1500 consécutifs de la séquence dont il est issu.

Par fragment représentatif, on entend en particulier une séquence nucléique codant pour un fragment biologiquement actif d'un polypeptide, tel que défini plus loin.

Par fragment représentatif, on entend également les séquences intergéniques, et en particulier les séquences nucléotidiques portant les signaux de régulation (promoteurs, terminateurs, voire enhanceurs, ...).

Parmi lesdits fragments représentatifs, on préfère ceux ayant des séquences nucléotidiques correspondant à des cadres ouverts de lecture, dénommés séquences ORFs (ORF pour « Open Reading Frame »), compris en général entre un codon d'initiation et un codon stop, ou entre deux codons stop, et codant pour des polypeptides, de préférence d'au moins 100 acides aminés, tel que par exemple, sans s'y limiter, les séquences ORFs qui seront décrites par la suite.

La numérotation des séquences nucléotidiques ORFs qui sera utilisée par la suite dans la présente description correspond à la numérotation des séquences d'acides

aminés des protéines codées par lesdites ORFs pour les peptides de séquence SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No.2346 à SEQ ID No.4481.

Les fragments représentatifs selon l'invention peuvent être obtenus par exemple par amplification spécifique telle que la PCR ou après digestion par des enzymes de restriction appropriés de séquences nucléotidiques selon l'invention, cette méthode étant
5 décrite en particulier dans l'ouvrage de Sambrook et al.. Lesdits fragments représentatifs peuvent également être obtenus par synthèse chimique lorsque leur taille n'est pas trop importante, selon des méthodes bien connues de l'homme du métier.

Parmi les séquences contenant des séquences de l'invention, ou des fragments
10 représentatifs, on entend également les séquences qui sont naturellement encadrées par des séquences qui présentent au moins 75 %, 80 %, 85 %, 90 %, 95 % ou 98 % d'identité avec les séquences selon l'invention.

Par séquence nucléotidique modifiée, on entend toute séquence nucléotidique obtenue par mutagenèse selon des techniques bien connues de l'homme du métier, et
15 comportant des modifications par rapport aux séquences normales, par exemple des mutations dans les séquences régulatrices et/ou promotrices de l'expression du polypeptide, notamment conduisant à une modification du taux d'expression ou de l'activité dudit polypeptide.

Par séquence nucléotidique modifiée, on entend également toute séquence
20 nucléotidique codant pour un polypeptide modifié tel que définit ci-après.

Concernant les séquences nucléiques ou ORF codant pour les peptides de séquence SEQ ID No. 2346 à SEQ ID No. 4481, ces séquences nucléiques ou ORF sont représentées respectivement par les séquences SEQ ID No. 4482 à SEQ ID No. 6617.

L'invention concerne avantageusement une séquence nucléotidique isolée de
25 *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi :

a) une séquence choisie parmi les séquences SEQ ID N° 6194,6236,5497,5791,5103,4705,5610,5234,4926,6331,6247,5842,5741,4921,5090, 5180,4706,4708,5677,6246,6411,5578,6446,6447,5607,6209,6215,5406,5658,4965, de préférence parmi les séquences SEQ ID N°4926,6331,5491,5234,6246,5842 ;

30 b) une séquence nucléotidique comportant au moins 75 % d'identité avec une séquence nucléotidique du a) ;

c) une séquence nucléotidique s'hybridant dans des conditions de forte stringence avec une séquence nucléotidique du a) ou b) et comprenant au moins 20 nucléotides ;

d) une séquence nucléotidique complémentaire ou d'ARN correspondant à une séquence telle que définie en a), b) ou c) ;

e) une séquence nucléotidique d'un fragment représentatif d'une séquence telle que définie en a), b), c) ou d) et comprenant au moins 20 nucléotides ; et

5 f) une séquence telle que définie en a), b), c), d) ou e) modifiée et comportant au plus 10 % de nucléotides modifiés par rapport à la séquence de référence ;

et en ce qu'elle code pour une protéine de surface avec un motif d'ancrage LPXTG.

L'invention concerne également les polypeptides codés par ces séquences.

10 L'invention concerne également avantageusement une séquence nucléotidique isolée de *Streptococcus agalactiae* caractérisée en ce qu'elle est choisie parmi les séquences

	SEQ	ID
	N°6035,6137,6335,6377,6386,4495,4596,4636,4730,4816,4836,4906,4920,4925,5158,	
	5247, 5306,5417,5450,5486,5559,5591,5677,5732,5799,5800,5861,5923 ;	

15 qu'elle code pour une lipoprotéine. L'invention concerne également les polypeptides codés par ces séquences.

L'invention concerne également avantageusement une séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi les séquences

	SEQ	ID
20	N°4861,6214,6061,6517,6518,6519,4743,6343,6342,5326,4952,5619,5618,5617,5616,	
	5615,5614,5613,5611,5696,5971,5233,5602,5156,5574,5573,5654,5656,5526,5527,552	
	9,5534,5625,5626,6223,6229,6230,6231,6232,6233,5764,6095,5089,5466,5465 ;	

et en ce qu'elle code pour une protéine impliquée dans la biosynthèse de composés polysaccharidiques de paroi. L'invention concerne également les polypeptides codés par

25 ces séquences.

Les fragments représentatifs selon l'invention peuvent également être des sondes ou amorces, qui peuvent être utilisées dans des procédés de détection, d'identification, de dosage ou d'amplification de séquences nucléiques.

30 Une sonde ou amorce se définit, au sens de l'invention, comme étant un fragment d'acides nucléiques simple brin ou un fragment double brin dénaturé comprenant par exemple de 12 bases à quelques kb, notamment de 15 à quelques centaines de bases, de préférence de 15 à 50 ou 100 bases, et possédant une spécificité d'hybridation dans des conditions déterminées pour former un complexe d'hybridation avec un acide nucléique cible.

Les sondes et amorces selon l'invention peuvent être marquées directement ou indirectement par un composé radioactif ou non radioactif par des méthodes bien connues de l'homme du métier, afin d'obtenir un signal détectable et/ou quantifiable (brevet FR 78 10975 et bDNA de Chiron EP 225 807 et EP 510 085).

5 Les séquences non marquées de polynucléotides selon l'invention peuvent être utilisées directement comme sonde ou amorce.

Les séquences sont généralement marquées pour obtenir des séquences utilisables pour de nombreuses applications. Le marquage des amorces ou des sondes selon l'invention est réalisé par des éléments radioactifs ou par des molécules non
10 radioactives.

Parmi les isotopes radioactifs utilisés, on peut citer le ^{32}P , le ^{33}P , le ^{35}S , le ^3H ou le ^{125}I . Les entités non radioactives sont sélectionnées parmi les ligands tels la biotine, l'avidine, la streptavidine, la dioxygénine, les haptènes, les colorants, les agents luminescents tels que les agents radioluminescents, chémoluminescents,
15 bioluminescents, fluorescents, phosphorescents.

Les polynucléotides selon l'invention peuvent ainsi être utilisés comme amorce et/ou sonde dans des procédés mettant en oeuvre notamment la technique de PCR (amplification en chaîne par polymérase) (Rolfs et al., 1991, Berlin : Springer-Verlag). Cette technique nécessite le choix de paires d'amorces oligonucléotidiques encadrant le
20 fragment qui doit être amplifié. On peut, par exemple, se référer à la technique décrite dans le brevet américain U.S. N° 4,683,202. Les fragments amplifiés peuvent être identifiés, par exemple après une électrophorèse en gel d'agarose ou de polyacrylamide, ou après une technique chromatographique comme la filtration sur gel ou la chromatographie échangeuse d'ions, puis séquencés. La spécificité de l'amplification
25 peut être contrôlée en utilisant les séquences nucléotidiques de polynucléotides de l'invention comme matrice, des plasmides contenant ces séquences ou encore les produits d'amplification dérivés. Les fragments nucléotidiques amplifiés peuvent être utilisés comme réactifs dans des réactions d'hybridation afin de mettre en évidence la présence, dans un échantillon biologique, d'un acide nucléique cible de séquence
30 complémentaire à celle desdits fragments nucléotidiques amplifiés.

L'invention vise également les acides nucléiques susceptibles d'être obtenus par amplification à l'aide d'amorces selon l'invention.

D'autres techniques d'amplification de l'acide nucléique cible peuvent être avantageusement employées comme alternative à la PCR (PCR-like) à l'aide de couple

d'amorces de séquences nucléotidiques selon l'invention. Par PCR-like on entend désigner toutes les méthodes mettant en œuvre des reproductions directes ou indirectes des séquences d'acides nucléiques, ou bien dans lesquelles les systèmes de marquage ont été amplifiés, ces techniques sont bien entendu connues. En général il s'agit de

5 l'amplification de l'ADN par une polymérase ; lorsque l'échantillon d'origine est un ARN il convient préalablement d'effectuer une transcription reverse. Il existe actuellement de très nombreux procédés permettant cette amplification, comme par exemple la technique SDA (Strand Displacement Amplification) ou technique d'amplification à déplacement de brin (Walker et al., 1992, Nucleic Acids Res.

10 20:1691), la technique TAS (Transcription-based Amplification System) décrite par Kwoh et al. (1989, Proc. Natl. Acad. Sci., USA, 86, 1173), la technique 3SR (Self-Sustained Sequence Replication) décrite par Guatelli et al. (1990, Proc. Natl. Acad. Sci., USA 87:1874), la technique NASBA (Nucleic Acid Sequence Based Amplification) décrite par Kievitis et al. (1991, J. Virol. Methods, 35, 273), la technique TMA

15 (Transcription Mediated Amplification), la technique LCR (Ligase Chain Reaction) décrite par Landegren et al. (1988, Science 241, 1077), la technique de RCR (Repair Chain Reaction) décrite par Segev (1992, Kessler C. Springer Verlag, Berlin, New-York, 197-205), la technique CPR (Cycling Probe Reaction) décrite par Duck et al. (1990, Biotechniques, 9, 142), la technique d'amplification à la Q-béta-réplacase décrite

20 par Miele et al. (1983, J. Mol. Biol., 171, 281). Certaines de ces techniques ont depuis été perfectionnées.

Dans le cas où le polynucléotide cible à détecter est un ARNm, on utilise avantageusement, préalablement à la mise en oeuvre d'une réaction d'amplification à l'aide des amorces selon l'invention ou à la mise en œuvre d'un procédé de détection à

25 l'aide des sondes de l'invention, une enzyme de type transcriptase inverse afin d'obtenir un ADNc à partir de l'ARNm contenu dans l'échantillon biologique. L'ADNc obtenu servira alors de cible pour les amorces ou les sondes mises en oeuvre dans le procédé d'amplification ou de détection selon l'invention.

La technique d'hybridation de sondes peut être réalisée de manières diverses

30 (Matthews et al., 1988, Anal. Biochem., 169, 1-25). La méthode la plus générale consiste à immobiliser l'acide nucléique extrait des cellules de différents tissus ou de cellules en culture sur un support (tels que la nitrocellulose, le nylon, le polystyrène) et à incuber, dans des conditions bien définies, l'acide nucléique cible immobilisé avec la sonde. Après l'hybridation, l'excès de sonde est éliminé et les molécules hybrides

formées sont détectées par la méthode appropriée (mesure de la radioactivité, de la fluorescence ou de l'activité enzymatique liée à la sonde).

5 Selon un autre mode de mise en œuvre des sondes nucléiques selon l'invention, ces dernières peuvent être utilisées comme sondes de capture. Dans ce cas, une sonde, dite « sonde de capture », est immobilisée sur un support et sert à capturer par hybridation spécifique l'acide nucléique cible obtenu à partir de l'échantillon biologique à tester et l'acide nucléique cible est ensuite détecté grâce à une seconde sonde, dite « sonde de détection », marquée par un élément facilement détectable.

10 Parmi les fragments d'acides nucléiques intéressants, il faut ainsi citer en particulier les oligonucléotides anti-sens, c'est-à-dire dont la structure assure, par hybridation avec la séquence cible, une inhibition de l'expression du produit correspondant. Il faut également citer les oligonucléotides sens qui, par interaction avec des protéines impliquées dans la régulation de l'expression du produit correspondant, induiront soit une inhibition, soit une activation de cette expression.

15 De façon préférée, les sondes ou amorces selon l'invention sont immobilisées sur un support, de manière covalente ou non covalente. En particulier, le support peut être une puce à ADN ou un filtre à haute ou moyenne densité, également objets de la présente invention (brevets WO 97/29212, WO 98/27317, WO 97/10365 et WO 92/10588).

20 On entend désigner par puce à ADN ou filtre haute densité, un support sur lequel sont fixées des séquences d'ADN, chacune d'entre elles pouvant être repérée par sa localisation géographique. Ces puces ou filtres diffèrent principalement par leur taille, le matériau du support, et éventuellement le nombre de séquences d'ADN qui y sont fixées.

25 On peut fixer les sondes ou amorces selon la première invention sur des supports solides, en particulier les puces à ADN, par différents procédés de fabrication. En particulier, on peut effectuer une synthèse *in situ* par adressage photochimique ou par jet d'encre. D'autres techniques consistent à effectuer une synthèse *ex situ* et à fixer les sondes sur le support de la puce à ADN par adressage mécanique, électronique ou par jet d'encre. Ces différents procédés sont bien connus de l'homme du métier.

30 Une séquence nucléotidique (sonde ou amorce) selon l'invention permet donc la détection et/ou l'amplification de séquences nucléiques spécifiques. En particulier, la détection de cesdites séquences est facilitée lorsque la sonde est fixée sur une puce à ADN, ou à un filtre haute densité.

L'utilisation de puces à ADN ou de filtres à haute densité permet en effet de déterminer l'expression de gènes dans un organisme présentant une séquence génomique proche de *Streptococcus agalactiae* et le typage de la souche en cause.

La séquence génomique de *Streptococcus agalactiae*, complétée par
5 l'identification des gènes de ces organismes, telle que présentée dans la présente invention, sert de base à la construction de ces puces à ADN ou filtre.

La préparation de ces filtres ou puces consiste à synthétiser des oligonucléotides, correspondant aux extrémités 5' et 3' des gènes ou à des fragments plus internes pour amplifier des fragments d'une taille adaptée, par exemple comprise environ entre 300 et
10 800 bases. Ces oligonucléotides sont choisis en utilisant la séquence génomique et ses annotations divulguées par la présente invention. La température d'appariement des ces oligonucléotides aux places correspondantes sur l'ADN doit être approximativement la même pour chaque oligonucleotide. Ceci permet de préparer des fragments d'ADN correspondant à chaque gène par l'utilisation de conditions de PCR appropriées dans un
15 environnement hautement automatisé. Les fragments amplifiés sont ensuite immobilisés sur des filtres ou des supports en verre, silicium ou polymères synthétiques et ces milieux sont utilisés pour l'hybridation.

La disponibilité de tels filtres et/ou puces et de la séquence génomique correspondante annotée permet d'étudier l'expression de grands ensembles, voire de la
20 totalité des gènes dans les micro-organismes associés à *Streptococcus agalactiae* et *Streptococcus agalactiae* CIP 82.45 (ATCC 12403), en préparant les ADN compleméntaires, et en les hybridant à l'ADN ou aux oligonucléotides immobilisés sur les filtres ou les puces. De même, les filtres et/ou les puces permettent d'étudier la variabilité des souches ou des espèces, en préparant l'ADN de ces organismes et en les
25 hybridant à l'ADN ou aux oligonucléotides immobilisés sur les filtres ou les puces.

Les différences entre les séquences génomiques des différentes souches ou espèces peuvent grandement affecter l'intensité de l'hybridation et, par conséquent, perturber l'interprétation des résultats. Il peut donc être nécessaire d'avoir la séquence précise des gènes de la souche que l'on souhaite étudier. La méthode de détection des
30 gènes décrite plus loin en détail, impliquant la détermination de la séquence de fragments aléatoires d'un génome, et les organisant d'après la séquence du génome de *Streptococcus agalactiae*, notamment de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) divulguée dans la présente invention, peut être très utile.

Les séquences nucléotidiques selon l'invention peuvent être utilisées dans des puces à ADN pour effectuer l'analyse de mutations. Cette analyse repose sur la constitution de puces capables d'analyser chaque base d'une séquence nucléotidique selon l'invention. On pourra notamment à cette fin mettre en œuvre les techniques de microséquençage sur puce à ADN. Les mutations sont détectées par extension d'amorces immobilisées hybridant à la matrice des séquences analysées, juste en position adjacente de celle du nucléotide muté recherché. Une matrice simple brin, ARN ou ADN, des séquences à analyser sera avantageusement préparée selon des méthodes classiques, à partir de produits amplifiés selon les techniques de type PCR. Les matrices d'ADN simple brin, ou d'ARN ainsi obtenues sont alors déposées sur la puce à ADN, dans des conditions permettant leur hybridation spécifique aux amorces immobilisées. Une polymérase thermostable, par exemple la Tth ou la Taq ADN polymérase, étend spécifiquement l'extrémité 3' de l'amorce immobilisée avec un analogue de nucléotide marqué complémentaire du nucléotide en position du site variable ; par exemple, un cyclage thermique est réalisé en présence des didéoxyribonucléotides fluorescents. Les conditions expérimentales seront adaptées notamment aux puces employées, aux amorces immobilisées, aux polymérases employées, et au système de marquage choisi. Un avantage du microséquençage, par rapport aux techniques basées sur l'hybridation de sondes, est qu'il permet d'identifier tous les nucléotides variables avec une discrimination optimale dans des conditions de réactions homogènes ; utilisé sur des puces à ADN, il permet une résolution et une spécificité optimales pour la détection routinière et industrielle de mutations en multiplex.

Une puce à ADN ou un filtre peut être un outil extrêmement intéressant pour la détermination, la détection et/ou l'identification d'un micro-organisme. Ainsi, on préfère également les puces à ADN selon l'invention qui contiennent en outre au moins une séquence nucléotidique d'un micro-organisme autre que *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou *Streptococcus agalactiae*, immobilisée sur le support de ladite puce. De préférence, le micro-organisme choisi l'est parmi les bactéries du genre *Streptococcus* (ci-après désignées comme bactéries associées à *Streptococcus agalactiae*), ou les variants de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403).

Une puce à ADN ou un filtre selon l'invention est un élément très utile de certains kits ou nécessaires pour la détection et/ou l'identification de micro-organismes, en particulier les bactéries appartenant à l'espèce *Streptococcus agalactiae* ou les micro-organismes associés, également objets de l'invention.

Par ailleurs, les puces à ADN ou les filtres selon l'invention, contenant des sondes ou amorces spécifiques de *Streptococcus agalactiae*, sont des éléments très avantageux de kits ou nécessaires pour la détection et/ou la quantification de l'expression de gènes de *Streptococcus agalactiae* (ou de micro-organismes associés).

5 En effet, le contrôle de l'expression des gènes est un point critique pour optimiser la croissance et le rendement d'une souche, soit en permettant l'expression d'un ou plusieurs gènes nouveaux, soit en modifiant l'expression de gènes déjà présents dans la cellule. La présente invention fournit l'ensemble des séquences naturellement actives chez *Streptococcus agalactiae* permettant l'expression des gènes. Elle permet
10 ainsi la détermination de l'ensemble des séquences exprimées chez *Streptococcus agalactiae*. Elle fournit également un outil permettant de repérer les gènes dont l'expression suit un schéma donné. Pour réaliser cela, l'ADN de tout ou partie des gènes de *Streptococcus agalactiae* peut être amplifié grâce à des amorces selon l'invention, puis fixé à un support comme par exemple le verre ou le nylon ou une puce à ADN, afin
15 de construire un outil permettant de suivre le profil d'expression de ces gènes. Cet outil, constitué de ce support contenant les séquences codantes sert de matrice d'hybridation à un mélange de molécules marquées reflétant les ARNs messagers exprimés dans la cellule (en particulier les sondes marquées selon l'invention). En répétant cette expérience à différents instants et en combinant l'ensemble de ces données par un
20 traitement approprié, on obtient alors les profils d'expression de l'ensemble de ces gènes. La connaissance des séquences qui suivent un schéma de régulation donné peut aussi être mise à profit pour rechercher de manière dirigée, par exemple par homologie, d'autres séquences suivant globalement, mais de manière légèrement différente le même schéma de régulation. En complément, il est possible d'isoler chaque séquence de
25 contrôle présente en amont des segments servant de sondes et d'en suivre l'activité à l'aide de moyen approprié comme un gène rapporteur (luciférase, β -galactosidase, GFP). Ces séquences isolées peuvent ensuite être modifiées et assemblées par ingénierie métabolique avec des séquences d'intérêt en vue de leur expression optimale.

30 L'invention concerne également les polypeptides codés par une séquence nucléotidique selon l'invention, de préférence, par un fragment représentatif des séquences précédentes et correspondant à une séquence ORF. En particulier, les polypeptides de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) de SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481 sont objet de l'invention.

L'invention comprend également les polypeptides caractérisés en ce qu'ils comprennent un polypeptide choisi parmi :

- a) un polypeptide selon l'invention ;
- b) un polypeptide présentant au moins 80 % de préférence 85 %, 90 %, 95 % et 98 %
5 d'identité avec un polypeptide selon l'invention ;
- c) un fragment d'au moins 5 acides aminés, de préférence d'au moins 10, 15, 20, 25, 30, 40, 50, 75 et 100 acides aminés d'un polypeptide selon l'invention, ou tel que défini en b) ;
- d) un fragment biologiquement actif d'un polypeptide selon l'invention, ou tel que
10 défini en b) ou c) ; et
- e) un polypeptide selon l'invention, ou tel que défini en b), c) ou d) modifié et comportant au plus 10 %, 5 % ou 1 % d'acides aminés modifiés par rapport à la séquence de référence.

Les séquences nucléotidiques codant pour les polypeptides décrits
15 précédemment sont également objet de l'invention.

Dans la présente description, les termes polypeptides, séquences polypeptidiques, peptides et protéines sont interchangeables. Le terme polypeptide comprend toute séquence d'acides aminés permettant de générer une réponse anticorps.

Il doit être compris que l'invention ne concerne pas les polypeptides sous forme
20 naturelle, c'est-à-dire qu'ils ne sont pas pris dans leur environnement naturel. En revanche, elle concerne ceux qui ont pu être isolés ou obtenus par purification à partir de sources naturelles, ou bien obtenus par recombinaison génétique, ou par synthèse chimique, et qu'ils peuvent alors comporter des acides aminés non naturels comme cela sera décrit plus loin.

25 Par polypeptide présentant un certain pourcentage d'identité avec un autre, que l'on désignera également par polypeptide homologue, on entend désigner les polypeptides présentant par rapport aux polypeptides naturels, certaines modifications, en particulier une délétion, addition ou substitution d'au moins un acide aminé, une troncation, un allongement, une solution chimérique et/ou une mutation, ou les
30 polypeptides présentant des modifications post-traductionnelles. Parmi les polypeptides homologues, on préfère ceux dont la séquence d'acides aminés présentent au moins 80 %, de préférence 85 %, 90 %, 95 % et 98 % d'homologie avec les séquences d'acides aminés des polypeptides selon l'invention. Dans le cas d'une substitution, un ou plusieurs acide(s) aminé(s) consécutif(s) ou non consécutif(s) sont remplacés par des

acides aminés « équivalents ». L'expression « acides aminés équivalents » vise ici à désigner tout acide aminé susceptible d'être substitué à l'un des acides aminés de la structure de base sans cependant modifier essentiellement les activités biologiques des peptides correspondant telles qu'elles seront définies par la suite.

- 5 Ces acides aminés équivalents peuvent être déterminés soit en s'appuyant sur leur homologie de structure avec les acides aminés auxquels ils se substituent, soit sur des résultats d'essais comparatifs d'activité biologique entre les différents polypeptides susceptibles d'être effectués.

10 A titre d'exemple, on mentionne les possibilités de substitution susceptibles d'être effectuées sans qu'il résulte en une modification approfondie de l'activité biologique du polypeptide modifié correspondant. On peut remplacer ainsi la leucine par la valine ou l'isoleucine, l'acide aspartique par l'acide glutamine, la glutamine par l'asparagine, l'arginine par la lysine, etc., les substitutions inverses étant naturellement envisageables dans les mêmes conditions.

- 15 Les polypeptides homologues correspondent également aux polypeptides codés par les séquences nucléotidiques homologues ou identiques, telles que définies précédemment et comprennent ainsi dans la présente définition des polypeptides mutés ou correspondant à des variations inter ou intra espèces, pouvant exister chez *Streptococcus*, et qui correspondent notamment à des troncatures, substitutions,
- 20 délétions et/ou additions, d'au moins un résidu d'acides aminés.

 Il est entendu que l'on calcule le pourcentage d'identité entre deux polypeptides de la même façon qu'entre deux séquences d'acides nucléiques. Ainsi, le pourcentage d'identité entre deux polypeptides est calculé après alignement optimal de ces deux séquences, sur une fenêtre d'homologie maximale. Pour définir ladite fenêtre

25 d'homologie maximale, on peut utiliser les mêmes algorithmes que pour les séquences d'acide nucléique.

 Par fragment biologiquement actif d'un polypeptide selon l'invention, on entend désigner en particulier un fragment de polypeptide, tel que défini ci-après, présentant au moins une des caractéristiques biologiques des polypeptides selon l'invention,

30 notamment en ce qu'il est capable d'exercer de manière générale une activité même partielle, telle que par exemple :

 - une activité enzymatique (métabolique) ou une activité pouvant être impliquée dans la biosynthèse ou la biodégradation de composés organiques ou inorganiques ;

- une activité structurale (enveloppe cellulaire, molécule chaperonne, ribosome) ;
- une activité de transport (d'énergie, d'ion) ; ou dans la sécrétion de protéine ;
- une activité dans le processus de réplication, amplification, préparation, transcription, traduction ou maturation, notamment de l'ADN, de l'ARN ou des protéines.

Par fragment de polypeptide selon l'invention, on entend désigner un polypeptide comportant au minimum 5 acides aminés, de préférence d'au moins 10, 15, 20, 25, 30, 40, 50, 75, 100 et 150 acides aminés.

Les fragments de polypeptides peuvent correspondre à des fragments isolés ou purifiés naturellement présents dans les souches de *Streptococcus*, ou à des fragments qui peuvent être obtenus par clivage dudit polypeptide par une enzyme protéolytique telle que la trypsine ou la chymotrypsine ou la collagénase, par un réactif chimique (bromure de cyanogène, CNBr) ou en plaçant ledit polypeptide dans un environnement très acide (par exemple à pH = 2,5). Des fragments polypeptidiques peuvent également être préparés par synthèse chimique, à partir d'hôtes transformés par un vecteur d'expression selon l'invention qui contiennent un acide nucléique permettant l'expression dudit fragment, et placé sous le contrôle des éléments de régulation et/ou d'expression appropriés.

Par « polypeptide modifié » d'un polypeptide selon l'invention, on entend désigner un polypeptide obtenu par recombinaison génétique ou par synthèse chimique comme décrit plus loin, qui présente au moins une modification par rapport à la séquence normale. Ces modifications peuvent être notamment portées sur des acides aminés nécessaires pour la spécificité ou l'efficacité de l'activité, ou à l'origine de la conformation structurale, de la charge, ou de l'hydrophobicité du polypeptide selon l'invention. On peut ainsi créer des polypeptides d'activité équivalente, augmentée ou diminuée, ou de spécificité équivalente, plus étroite ou plus large. Parmi les polypeptides modifiés, il faut citer les polypeptides dans lesquels jusqu'à cinq acides aminés peuvent être modifiés, tronqués à l'extrémité N ou C-terminale, ou bien délétés, ou ajoutés.

Comme cela est indiqué, les modifications d'un polypeptide ont pour objectif notamment :

- de permettre sa mise en œuvre dans des procédés de biosynthèse ou de biodégradation de composés organiques ou inorganiques,

- de permettre sa mise en œuvre dans des procédés de réplication, d'amplification, de réparation et règle de transcription, de traduction, ou de maturation notamment de l'ADN, l'ARN, ou de protéines,

- de permettre sa sécrétion améliorée,
- 5 - de modifier sa solubilité, l'efficacité ou la spécificité de son activité, ou encore de faciliter sa purification.

La synthèse chimique présente également l'avantage de pouvoir utiliser des acides aminés non naturels ou des liaisons non peptidiques. Ainsi, il peut être intéressant d'utiliser des acides aminés non naturels, par exemple sous forme D, ou des
10 analogues d'acides aminés, notamment des formes souffrées.

La présente invention fournit la séquence nucléotidique du génome de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) sous forme de contigs, ainsi que certaines séquences polypeptidiques.

D'une manière préférée, l'invention est relative à une séquence nucléotidique
15 selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la biosynthèse des acides aminés.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus*
20 *agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la biosynthèse des cofacteurs, groupes prosthétiques et transporteurs.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide d'enveloppe cellulaire ou présent à la surface de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou pour
25 un de ses fragments.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la machinerie cellulaire.

30 De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme intermédiaire central.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme énergétique.

- 5 De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme des acides gras et des phospholipides.

- De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme des nucléotides, des purines, des pyrimidines ou nucléosides.
- 10

- De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans les fonctions de régulation.
- 15

- De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de réplication.
- 20

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de transcription.

- 25 De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de traduction.

- De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de transport et de liaison des protéines.
- 30

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus*

agalactiae CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans l'adaptation aux conditions atypiques.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus*
5 *agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments dans la sensibilité aux médicaments et analogues.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus*
10 *agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans les fonctions relatives aux transposons.

De manière préférée, l'invention est relative à une séquence nucléotidique selon l'invention, caractérisée en ce qu'elle code pour un polypeptide spécifique de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments.

Sous un autre aspect, de manière préférée, l'invention a pour objet un
15 polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la biosynthèse des acides aminés.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de
20 *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la biosynthèse des cofacteurs, groupes prosthétiques et transporteurs.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide d'enveloppe cellulaire ou de surface de *Streptococcus agalactiae* CIP 82.45 (ATCC
25 12403) ou un de ses fragments.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans la machinerie cellulaire.

30 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme intermédiaire central.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme énergétique.

- 5 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme des acides gras et des phospholipides.

- 10 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le métabolisme des nucléotides, des purines, des pyrimidines ou nucléosides.

- 15 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans les fonctions de régulation.

- 20 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de réplication.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de transcription.

- 25 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de traduction.

- 30 Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans le processus de transport et de liaison des protéines.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de

Streptococcus agalactiae CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans l'adaptation aux conditions atypiques.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de
5 *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments dans la sensibilité aux médicaments et analogues.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide de
10 *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments impliqué dans les fonctions relatives aux transposons.

Sous un autre aspect, de manière préférée, l'invention a pour objet un polypeptide selon l'invention, caractérisé en ce qu'il s'agit d'un polypeptide spécifique de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un de ses fragments.

L'invention a également pour objet les opérons impliqués dans la synthèse
15 d'antibiotiques et/ou de toxines.

Le tableau 1 fournit la liste de certains polypeptides selon l'invention, ainsi que leur localisation dans les séquences SEQ ID No. 1 à SEQ ID No. 139, et les analogies observées après comparaison dans les bases de données.

20 Il est important de noter toutefois qu'un organisme vivant est un tout et doit être pris comme tel. Ainsi, afin de pouvoir se développer et exhiber ses propriétés, tout organisme a besoin d'interactions entre les différentes voies métaboliques. Ainsi, la classification énoncée ci-dessus ne doit pas être considérée comme limitative, un gène pouvant être impliqué dans deux voies métaboliques distinctes.

25 La présente invention a également pour objet les séquences nucléotidiques et/ou de polypeptides selon l'invention, caractérisées en ce que lesdites séquences sont enregistrées sur un support d'enregistrement dont la forme et la nature facilitent la lecture, l'analyse et/ou l'exploitation de ladite ou desdites séquence(s). Ces supports peuvent également contenir d'autres informations extraites de la présente invention,
30 notamment les analogies avec des séquences déjà connues, et/ou des informations concernant les séquences nucléotidiques et/ou de polypeptides d'autres micro-organismes afin de faciliter l'analyse comparative et l'exploitation des résultats obtenus.

Parmi cesdits supports d'enregistrement, on préfère en particulier les supports lisibles par un ordinateur, tels les supports magnétiques, optiques, électriques ou

hybrides, en particulier les disquettes informatiques, les CD-ROM, les serveurs informatiques. De tels supports d'enregistrement sont également objet de l'invention.

Les supports d'enregistrement selon l'invention, avec les informations apportées, sont très utiles pour le choix d'amorces ou de sondes nucléotidiques pour la
5 détermination de gènes dans *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou souches proches de cet organisme. De même, l'utilisation de ces supports pour l'étude du polymorphisme génétique de souches proches de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403), en particulier par la détermination des régions de colinéarité, est très utile dans la mesure où ces supports fournissent non seulement la séquence
10 nucléotidique du génome de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403), mais également l'organisation génomique dans ladite séquence. Ainsi, les utilisations de supports d'enregistrement selon l'invention sont également des objets de l'invention.

L'analyse d'homologie entre différentes séquences s'effectue en effet avantageusement à l'aide de logiciels de comparaison de séquences, tels le logiciel
15 Blast, ou les logiciels de la trousse GCG, décrits précédemment.

L'invention vise également les vecteurs de clonage et/ou d'expression, qui contiennent une séquence nucléotidique selon l'invention.

Les vecteurs selon l'invention comportent de préférence des éléments qui permettent l'expression et/ou la sécrétion des séquences nucléotidiques dans une cellule
20 hôte déterminée.

Le vecteur doit alors comporter un promoteur, des signaux d'initiation et de terminaison de la traduction, ainsi que des régions appropriées de régulation de la transcription. Il doit pouvoir être maintenu de façon stable dans la cellule hôte et peut éventuellement posséder des signaux particuliers qui spécifient la sécrétion de la
25 protéine traduite. Ces différents éléments sont choisis et optimisés par l'homme du métier en fonction de l'hôte cellulaire utilisé. A cet effet, les séquences nucléotidiques selon l'invention peuvent être insérées dans des vecteurs à réplication autonome au sein de l'hôte choisi, ou être des vecteurs intégratifs de l'hôte choisi.

De tels vecteurs sont préparés par des méthodes couramment utilisées par
30 l'homme du métier, et les clones résultant peuvent être introduits dans un hôte approprié par des méthodes standards, telles que la lipofection, l'électroporation, le choc thermique, ou des méthodes chimiques.

Les vecteurs selon l'invention sont par exemple des vecteurs d'origine plasmidique ou virale. Ils sont utiles pour transformer des cellules hôtes afin de cloner ou d'exprimer les séquences nucléotidiques selon l'invention.

5 L'invention comprend également les cellules hôtes transformées par un vecteur selon l'invention.

L'hôte cellulaire peut être choisi parmi des systèmes procaryotes ou eucaryotes, par exemple les cellules bactériennes mais également les cellules de levure ou les cellules animales, en particulier les cellules de mammifères. On peut également utiliser des cellules d'insectes ou des cellules de plantes. Les cellules hôtes préférées selon
10 l'invention sont en particulier les cellules procaryotes, de préférence les bactéries appartenant au genre *Streptococcus*, à l'espèce *Streptococcus agalactiae*, plus particulièrement *Streptococcus agalactiae* CIP 82.45 (ATCC 12403), ou les micro-organismes associés à l'espèce *Streptococcus agalactiae*.

L'invention concerne également les végétaux et les animaux, excepté l'homme,
15 qui comprennent une cellule transformée selon l'invention. Les cellules transformées selon l'invention sont utilisables dans des procédés de préparation de polypeptides recombinants selon l'invention. Les procédés de préparation d'un polypeptide selon l'invention sous forme recombinante, caractérisés en ce qu'ils mettent en œuvre un vecteur et/ou une cellule transformée par un vecteur selon l'invention sont eux-mêmes
20 compris dans la présente invention. De préférence, on cultive une cellule transformée par un vecteur selon l'invention dans des conditions qui permettent l'expression dudit polypeptide et on récupère ledit peptide recombinant.

Ainsi qu'il a été dit, l'hôte cellulaire peut être choisi parmi des systèmes procaryotes ou eucaryotes. En particulier, il est possible d'identifier des séquences
25 nucléotidiques selon l'invention, facilitant la sécrétion dans un tel système procaryote ou eucaryote. Un vecteur selon l'invention portant une telle séquence peut donc être avantageusement utilisé pour la production de protéines recombinantes, destinées à être sécrétées. En effet, la purification de ces protéines recombinantes d'intérêt sera facilitée par le fait qu'elles sont présentes dans le surnageant de la culture cellulaire plutôt qu'à
30 l'intérieur des cellules hôtes.

On peut également préparer les polypeptides selon l'invention par synthèse chimique. Un tel procédé de préparation est également un objet de l'invention. L'homme du métier connaît les procédés de synthèse chimique, par exemple les techniques mettant en œuvre des phases solides (voir notamment Steward et al., 1984,

Solid phase peptides synthesis, Pierce Chem. Company, Rockford, 111, 2ème éd., (1984)) ou des techniques utilisant des phases solides partielles, par condensation de fragments ou par une synthèse en solution classique. Les polypeptides obtenus par synthèse chimique et pouvant comporter des acides aminés non naturels correspondants
5 sont également compris dans l'invention.

L'invention est en outre relative à des polypeptides hybrides présentant au moins un polypeptide ou un de ses fragments selon l'invention, et une séquence d'un polypeptide susceptible d'induire une réponse immunitaire chez l'homme ou l'animal.

Avantageusement, le déterminant antigénique est tel qu'il est susceptible
10 d'induire une réponse humorale et/ou cellulaire.

Un tel déterminant pourra comprendre un polypeptide ou un de ses fragments selon l'invention sous forme glycosylée, utilisé en vue d'obtenir des compositions immunogènes susceptibles d'induire la synthèse d'anticorps dirigés contre des épitopes multiples. Lesdits polypeptides ou leurs fragments glycosylés font également partie de
15 l'invention.

Ces molécules hybrides peuvent être constituées en partie d'une molécule porteuse de polypeptides ou de leurs fragments selon l'invention, associée à une partie éventuellement immunogène, en particulier un épitope de la toxine diphtérique, la toxine tétanique, un antigène de surface du virus de l'hépatite B (brevet FR 79 21811),
20 l'antigène VP1 du virus de la poliomyélite ou toute autre toxine ou antigène viral ou bactérien.

Les procédés de synthèse des molécules hybrides englobent les méthodes utilisées en génie génétique pour construire des séquences nucléotidiques hybrides codant pour les séquences polypeptidiques recherchées. On pourra, par exemple, se
25 référer avantageusement à la technique d'obtention de gènes codant pour des protéines de fusion décrite par Minton en 1984.

Lesdites séquences nucléotidiques hybrides codant pour un polypeptide hybride ainsi que les polypeptides hybrides selon l'invention, caractérisés en ce qu'il s'agit de polypeptides recombinants obtenus par l'expression desdites séquences nucléotidiques
30 hybrides, font également partie de l'invention.

L'invention comprend également les vecteurs caractérisés en ce qu'ils contiennent une desdites séquences nucléotidiques hybrides. Les cellules hôtes transformées par lesdits vecteurs, les animaux transgéniques comprenant une desdites cellules transformées ainsi que les procédés de préparation de polypeptides

recombinants utilisant lesdits vecteurs, lesdites cellules transformées et/ou lesdits animaux transgéniques font également partie de l'invention.

Le couplage entre un polypeptide selon l'invention et un polypeptide immunogène peut être effectué par voie chimique, ou par voie biologique. Ainsi, selon
5 l'invention, il est possible d'introduire un ou plusieurs élément(s) de liaison, notamment des acides aminés pour faciliter les réactions de couplage entre le polypeptide selon l'invention, et le polypeptide immunostimulateur, le couplage covalent de l'antigène immunostimulateur pouvant être réalisé à l'extrémité N ou C-terminale du polypeptide selon l'invention. Les réactifs bifonctionnels permettant ce couplage sont déterminés en
10 fonction de l'extrémité choisie pour réaliser ce couplage, et les techniques de couplage sont bien connues de l'homme du métier.

Les conjugués issus d'un couplage de peptides peuvent être également préparés par recombinaison génétique. Le peptide hybride (conjugué) peut en effet être produit par des techniques d'ADN recombinant, par insertion ou addition à la séquence d'ADN
15 codant pour le polypeptide selon l'invention, d'une séquence codant pour le ou les peptide(s) antigène(s), immunogène(s) ou haptène(s). Ces techniques de préparation de peptides hybrides par recombinaison génétique sont bien connues de l'homme du métier (voir par exemple Makrides, 1996, *Microbiological Reviews* 60, 512-538).

De préférence, ledit polypeptide immunitaire est choisi dans le groupe des
20 peptides contenant les anatoxines, notamment le toxoïde diphtérique ou le toxoïde tétanique, les protéines dérivées du Streptocoque (comme la protéine de liaison à la séralbumine humaine), les protéines membranaires OMPA et les complexes de protéines de membranes externes, les vésicules de membranes externes ou les protéines de chocs thermiques.

Les polypeptides hybrides selon l'invention sont très utiles pour obtenir des
25 anticorps monoclonaux ou polyclonaux, capables de reconnaître spécifiquement les polypeptides selon l'invention. En effet, un polypeptide hybride selon l'invention permet la potentialisation de la réponse immunitaire, contre le polypeptide selon l'invention couplé à la molécule immunogène. De tels anticorps monoclonaux ou
30 polyclonaux, leurs fragments, ou les anticorps chimériques, reconnaissant les polypeptides selon l'invention, sont également objet de l'invention.

Les anticorps monoclonaux spécifiques peuvent être obtenus selon la méthode classique de culture d'hybridome décrite par Köhler et Milstein (1975, *Nature* 256, 495).

Les anticorps selon l'invention sont par exemple des anticorps chimériques, des anticorps humanisés, des fragments Fab, ou F(ab')². Ils peuvent également se présenter sous forme d'immunoconjugués ou d'anticorps marqués afin d'obtenir un signal détectable et/ou quantifiable.

5 Ainsi, les anticorps selon l'invention peuvent être employés dans un procédé pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes:

- a) mise en contact de l'échantillon biologique avec un anticorps selon l'invention ;
- 10 b) mise en évidence du complexe antigène-anticorps éventuellement formé.

Les anticorps selon la présente invention sont également utilisables afin de détecter une expression d'un gène de *Streptococcus agalactiae* ou de micro-organismes associés. En effet, la présence du produit d'expression d'un gène reconnu par un anticorps spécifique dudit produit d'expression peut être détectée par la présence d'un

15 complexe antigène-anticorps formé après la mise en contact de la souche de *Streptococcus agalactiae* ou du micro-organisme associé avec un anticorps selon l'invention. La souche bactérienne utilisée peut avoir été « préparée », c'est-à-dire centrifugée, lysée, placée dans un réactif approprié pour la constitution du milieu propice à la réaction immunologique. En particulier, on préfère un procédé de détection

20 de l'expression dans le gène, correspondant à un Western blot, pouvant être effectué après une électrophorèse sur gel de polyacrylamide d'un lysat de la souche bactérienne, en présence ou en l'absence de conditions réductrices (SDS-PAGE). Après migration et séparation des protéines sur le gel de polyacrylamide, on transfère lesdites protéines sur une membrane appropriée (par exemple en nylon) et on détecte la présence de la

25 protéine ou du polypeptide d'intérêt, par mise en contact de ladite membrane avec un anticorps selon l'invention.

Ainsi, la présente invention comprend également les kits ou nécessaires pour la mise en œuvre d'un procédé tel que décrit (de détection de l'expression d'un gène de *Streptococcus agalactiae* ou d'un micro-organisme associé, ou pour la détection et/ou

30 l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou un micro-organisme associé), comprenant les éléments suivants :

- a) un anticorps polyclonal ou monoclonal selon l'invention ;
- b) éventuellement, les réactifs pour la constitution du milieu propice à la réaction immunologique ;

c) éventuellement, les réactifs permettant la mise en évidence des complexes antigène-anticorps produits par la réaction immunologique.

Les polypeptides et les anticorps selon l'invention peuvent avantageusement être immobilisés sur un support, notamment une puce à protéines. Une telle puce à protéines
5 est un objet de l'invention, et peut également contenir au moins un polypeptide d'un micro-organisme autre que *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou un anticorps dirigé contre un composé d'un micro-organisme autre que *Streptococcus agalactiae* CIP 82.45 (ATCC 12403).

Les puces à protéines ou filtres à haute densité contenant des protéines selon
10 l'invention peuvent être construites de la même manière que les puces à ADN selon l'invention. En pratique, on peut effectuer la synthèse des polypeptides directement fixés sur la puce à protéines, ou effectuer une synthèse *ex situ* suivie d'une étape de fixation sur ladite puce du polypeptide synthétisé. Cette dernière méthode est préférable, lorsque l'on désire fixer des protéines de taille importante sur le support, ces protéines
15 étant avantageusement préparées par génie génétique. Toutefois, si l'on ne désire fixer que des peptides sur le support de ladite puce, il peut être plus intéressant de procéder à la synthèse desdits peptides directement *in situ*.

Les puces à protéines selon l'invention peuvent être avantageusement utilisées dans des kits ou nécessaires pour la détection et/ou l'identification de bactéries
20 associées à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme, ou de façon plus générale dans des kits ou nécessaires pour la détection et/ou l'identification de micro-organismes. Lorsque l'on fixe les polypeptides selon l'invention sur les puces à ADN, on recherche la présence d'anticorps dans les échantillons testés, la fixation d'un anticorps selon l'invention sur le support de la puce
25 à protéines permettant l'identification de la protéine dont ledit anticorps est spécifique.

De préférence, on fixe un anticorps selon l'invention sur le support de la puce à protéines, et on détecte la présence de l'antigène correspondant, spécifique de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou d'un micro-organisme associé.

Une puce à protéines ci-dessus décrite peut être utilisée pour la détection de
30 produits de gènes, pour établir un profil d'expression desdits gènes, en complément d'une puce à ADN selon l'invention.

Les puces à protéines selon l'invention sont également extrêmement utiles pour les expériences de protéomique, qui étudie les interactions entre les différentes protéines d'un micro-organisme donné. De façon simplifiée, on fixe des peptides représentatifs

des différentes protéines d'un organisme sur un support. Puis, on met ledit support en contact avec des protéines marquées, et après une étape optionnelle de rinçage, on détecte des interactions entre lesdites protéines marquées et les peptides fixés sur la puce à protéines.

- 5 Ainsi, les puces à protéines comprenant une séquence polypeptidique selon l'invention ou un anticorps selon l'invention sont objet de l'invention, ainsi que les kits ou nécessaires les contenant.

La présente invention couvre également un procédé de détection et/ou d'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45
10 (ATCC 12403) ou à un micro-organisme associé dans un échantillon biologique, qui met en œuvre une séquence nucléotidique selon l'invention.

Il doit être entendu que le terme échantillon biologique concerne dans la présente invention les échantillons prélevés à partir d'un organisme vivant (en particulier sang, tissus, organes ou autres prélevés à partir d'un mammifère) ou un
15 échantillon contenant du matériel biologique, c'est-à-dire de l'ADN ou de l'ARN. Un tel échantillon biologique comprend aussi les compositions alimentaires contenant des bactéries (par exemple les fromages, les produits laitiers), mais également des compositions alimentaires contenant des levures (bières, pains) ou autres. Le terme échantillon biologique concerne aussi les bactéries isolées à partir de ces prélèvements
20 ou compositions alimentaires.

Le procédé de détection et/ou d'identification mettant en œuvre les séquences nucléotidiques selon l'invention peut être de diverse nature.

On préfère un procédé comportant les étapes suivantes :

- a) éventuellement, isolement de l'ADN à partir de l'échantillon biologique à analyser,
25 ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ;
b) amplification spécifique de l'ADN de bactéries appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme associé à l'aide d'au moins une amorce selon l'invention ;
c) mise en évidence des produits d'amplification.

30 Ce procédé est basé sur l'amplification spécifique de l'ADN, en particulier par une réaction d'amplification en chaîne.

On préfère également un procédé comprenant les étapes suivantes :

- a) mise en contact d'une sonde nucléotidique selon l'invention avec un échantillon biologique, l'acide nucléique contenu dans l'échantillon biologique ayant, le cas

échéant, préalablement été rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à l'acide nucléique d'une bactérie appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme associé ;

- 5 b) mise en évidence de l'hybride éventuellement formé entre la sonde nucléotidique et l'ADN de l'échantillon biologique.

Un tel procédé ne doit pas être limité à la détection de la présence de l'ADN contenu dans l'échantillon biologique à tester, il peut être également mis en œuvre pour détecter l'ARN contenu dans ledit échantillon. Ce procédé englobe en particulier les

- 10 Southern et Northern blot.

Un autre procédé préféré selon l'invention comprend les étapes suivantes :

- a) mise en contact d'une sonde nucléotidique immobilisée sur un support selon l'invention avec un échantillon biologique, l'acide nucléique de l'échantillon, ayant, le cas échéant, été préalablement rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à l'acide nucléique d'une bactérie appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme associé ;
- 15 b) mise en contact de l'hybride formé entre la sonde nucléotidique immobilisée sur un support et l'acide nucléique contenu dans l'échantillon biologique, le cas échéant après élimination de l'ADN de l'échantillon biologique n'ayant pas hybridé avec la sonde, avec une sonde nucléotidique marquée selon l'invention ;
- 20 c) mise en évidence du nouvel hybride formé à l'étape b).

Ce procédé est avantageusement utilisé avec une puce à ADN selon l'invention, l'acide nucléique recherché s'hybridant avec une sonde présente à la surface de ladite puce, et étant détecté par l'utilisation d'une sonde marquée. Ce procédé est avantageusement mis en œuvre en combinant une étape préalable d'amplification de l'ADN ou de l'ADN complémentaire obtenu éventuellement par transcription inverse, à l'aide d'amorces selon l'invention.

- 25 Ainsi, la présente invention englobe également les kits ou nécessaires pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :
- 30 a) une sonde nucléotidique selon l'invention ;

- b) éventuellement, les réactifs nécessaires à la mise en œuvre d'une réaction d'hybridation ;
- c) éventuellement, au moins une amorce selon l'invention ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN.

5 De même, la présente invention englobe également les kits ou nécessaires pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :

- a) une sonde nucléotidique, dite sonde de capture, selon l'invention ;
- 10 b) une sonde oligonucléotidique, dite sonde de révélation, selon l'invention ;
- c) éventuellement, au moins une amorce selon l'invention ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN.

Enfin, les kits ou nécessaires pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) ou à un
15 micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :

- a) au moins une amorce selon l'invention ;
- b) éventuellement, les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN ;
- c) éventuellement, un composant permettant de vérifier la séquence du fragment
20 amplifié, plus particulièrement une sonde oligonucléotidique selon l'invention, sont également objet de la présente invention.

De préférence, lesdites amorces et/ou sondes et/ou polypeptides et/ou anticorps selon la présente invention utilisés dans les procédés et/ou kits ou nécessaires selon la présente invention sont choisis parmi les amorces et/ou sondes et/ou polypeptides et/ou
25 anticorps spécifiques de l'espèce *Streptococcus agalactiae* CIP 82.45 (ATCC 12403). De manière préférée, ces éléments sont choisis parmi les séquences nucléotidiques codant pour une protéine sécrétée, parmi les polypeptides sécrétés, ou parmi les anticorps dirigés contre des polypeptides sécrétés de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403).

30 La présente invention a également pour objet les souches de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) et/ou de micro-organismes associés contenant une ou plusieurs mutation(s) dans une séquence nucléotidique selon l'invention, en particulier une séquence ORF, ou leurs éléments régulateurs (en particulier promoteurs).

On préfère, selon la présente invention, les souches de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) présentant une ou plusieurs mutation(s) dans les séquences nucléotidiques codant pour des polypeptides impliqués dans la machinerie cellulaire, en particulier la sécrétion, le métabolisme intermédiaire central, le métabolisme énergétique, les processus de synthèse des acides aminés, de transcription et de traduction, de synthèse des polypeptides.

Lesdites mutations peuvent mener à une inactivation du gène, ou en particulier lorsqu'elles sont situées dans les éléments régulateurs dudit gène, à une surexpression de celui-ci.

L'invention concerne en outre l'utilisation d'une séquence nucléotidique selon l'invention, d'un polypeptide selon l'invention, d'un anticorps selon l'invention, d'une cellule selon l'invention, et/ou d'un animal transformé selon l'invention, pour la sélection de composé organique ou inorganique capable de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, et/ou de modifier la réplication cellulaire de cellules eucaryotes ou procaryotes ou capables d'induire, d'inhiber ou d'aggraver une pathologies liées à une infection par *Streptococcus agalactiae* ou un de ses micro-organismes associés.

L'invention comprend également une méthode de sélection de composés capables de se lier à un polypeptide ou un de ses fragments selon l'invention, capables de se lier à une séquence nucléotidique selon l'invention, ou capables de reconnaître un anticorps selon l'invention, et/ou capables de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, et/ou de modifier la croissance ou la réplication cellulaire de cellules eucaryotes ou procaryotes, ou capables d'induire, d'inhiber ou d'aggraver chez un organisme animal ou humain une pathologie liée à une infection par *Streptococcus*, par exemple par *Streptococcus agalactiae*, ou un de ses micro-organismes associés, caractérisée en ce qu'elle comprend les étapes suivantes :

- a) mise en contact dudit composé avec ledit polypeptide, ladite séquence nucléotidique, avec une cellule transformée selon l'invention et/ou administration dudit composé à un animal transformé selon l'invention ;
- b) détermination de la capacité dudit composé à se lier avec ledit polypeptide ou ladite séquence nucléotidique, ou de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, ou de moduler la croissance ou la réplication cellulaire, ou d'induire, d'inhiber ou d'aggraver chez ledit animal transformé les pathologies liées

à une infection par *Streptococcus*, par exemple *Streptococcus agalactiae* ou un de ses micro-organismes associés.

Les cellules et/ou les animaux transformés selon l'invention, pourront avantageusement servir de modèle et être utilisés dans des procédés pour étudier, identifier et/ou sélectionner des composés susceptibles d'être responsables de pathologies induites ou aggravées par *Streptococcus agalactiae*, ou susceptibles de prévenir et/ou de traiter ces pathologies. En particulier, les cellules hôtes transformées, notamment les bactéries de la famille des *Streptococcus* dont la transformation par un vecteur selon l'invention peut par exemple accroître ou inhiber son pouvoir infectieux, ou moduler les pathologies habituellement induites ou aggravées par l'infection, pourront être utilisées pour infecter des animaux dont on suivra l'apparition des pathologies. Ces animaux non transformés, infectés par exemple avec des bactéries *Streptococcus* transformées, pourront servir de modèle d'étude. De la même manière, les animaux transformés selon l'invention pourront être utilisés dans des procédés de sélection de composés susceptibles de prévenir et/ou de traiter les maladies dues à *Streptococcus*. Lesdits procédés utilisant lesdites cellules transformées et/ou animaux transformés, font partie de l'invention.

Les composés susceptibles d'être sélectionnés peuvent être des composés organiques tels que des polypeptides ou hydrates de carbone ou tous autres composés organiques ou inorganiques déjà connus, ou des composés organiques nouveaux élaborés à partir de techniques de modélisation moléculaire et obtenus par synthèse chimique ou biochimique, ces techniques étant connues de l'homme de l'art.

Lesdits composés sélectionnés pourront être utilisés pour moduler la croissance et/ou la réplication cellulaire de *Streptococcus agalactiae* ou tout autre micro-organisme associé et ainsi pour contrôler l'infection par ces micro-organismes. Lesdits composés selon l'invention pourront également être utilisés pour moduler la croissance et/ou la réplication cellulaire de toutes cellules eucaryotes ou procaryotes, notamment les cellules tumorales et les micro-organismes infectieux, pour lesquelles lesdits composés s'avéreront actifs, les méthodes permettant de déterminer lesdites modulations étant bien connues de l'homme de l'art.

On entend désigner par composé capable de moduler la croissance d'un micro-organisme tout composé permettant d'intervenir, de modifier, de limiter et/ou de réduire le développement, la croissance, la vitesse de prolifération et/ou la viabilité dudit micro-organisme.

Cette modulation peut être réalisée par exemple par un agent capable de se lier à une protéine et ainsi d'inhiber ou de potentialiser son activité biologique, ou capable de se lier à une protéine membranaire de la surface extérieure d'un micro-organisme et de bloquer la pénétration dudit micro-organisme dans la cellule hôte ou de favoriser l'action du système immunitaire de l'organisme infecté dirigé à l'encontre dudit micro-organisme. Cette modulation peut être également réalisée par un agent capable de se lier à une séquence nucléotidique d'un ADN ou ARN d'un micro-organisme et de bloquer par exemple l'expression d'un polypeptide dont l'activité biologique ou structurale est nécessaire à la croissance ou à la reproduction dudit micro-organisme.

On entend désigner par micro-organisme associé dans la présente invention, tout micro-organisme dont l'expression de gène peut être modulée, régulée, induite ou inhibée, ou dont la croissance ou la réplication cellulaire peut être également modulée par un composé de l'invention. On entend désigner également par micro-organisme associé dans la présente invention, tout micro-organisme comportant des séquences nucléotidiques ou des polypeptides selon l'invention. Ces micro-organismes peuvent dans certains cas comporter des polypeptides ou des séquences nucléotidiques identiques ou homologues à celles de l'invention et pourront également être détectés et/ou identifiés par les procédés ou kit de détection et/ou d'identification selon l'invention et également servir de cible pour les composés de l'invention. On entend aussi désigner par micro-organisme tout micro-organisme *Streptococcus agalactiae* de tout sérotype.

L'invention concerne les composés susceptibles d'être sélectionnés par une méthode de sélection selon l'invention.

L'invention concerne également une composition pharmaceutique comprenant un composé choisi parmi les composés suivants :

- a) une séquence nucléotidique selon l'invention ;
- b) un polypeptide selon l'invention ;
- c) un vecteur selon l'invention ;
- d) un anticorps selon l'invention ; et
- e) un composé susceptible d'être sélectionné par une méthode de sélection selon l'invention,

éventuellement en association avec un véhicule pharmaceutiquement acceptable.

La présente invention concerne en outre une composition pharmaceutique selon l'invention pour la prévention et le traitement d'une infection par une bactérie appartenant à l'espèce *Streptococcus agalactiae*.

La présente invention concerne en outre une composition pharmaceutique selon
5 l'invention, caractérisée en ce qu'elle comprend des anticorps dirigés contre les polypeptides spécifiques de *Streptococcus agalactiae*.

On entend désigner par quantité efficace, une quantité suffisante dudit composé ou anticorps, ou de polypeptide de l'invention, permettant de moduler la croissance de *Streptococcus agalactiae* ou d'un micro-organisme associé.

10 L'invention concerne aussi une composition pharmaceutique selon l'invention pour la prévention ou le traitement d'une infection par une bactérie appartenant au genre *Streptococcus* ou par un micro-organisme associé.

L'invention vise en outre une composition immunogène et/ou vaccinale, caractérisée en ce qu'elle comprend un ou plusieurs polypeptides selon l'invention et/ou
15 un ou plusieurs polypeptides hybrides selon l'invention.

L'invention comprend aussi l'utilisation d'une cellule transformée selon l'invention, pour la préparation d'une composition vaccinale.

L'invention vise également une composition vaccinale, caractérisée en ce qu'elle contient une séquence nucléotidique selon l'invention, un vecteur selon l'invention
20 et/ou une cellule transformée selon l'invention.

L'invention concerne en outre une composition immunogène capable d'induire une réponse immunitaire cellulaire ou humorale pour la prévention ou le traitement d'une infection par une bactérie appartenant à l'espèce *Streptococcus agalactiae*, caractérisée en ce qu'elle comprend une composition immunogène ou une composition
25 vaccinale selon l'invention, en association avec un véhicule pharmaceutiquement acceptable et éventuellement un ou plusieurs adjuvants de l'immunité appropriés.

L'invention concerne également les compositions vaccinales selon l'invention, pour la prévention ou le traitement d'une infection par une bactérie appartenant au genre *Streptococcus* ou par un micro-organisme associé.

30 De manière préférée, les compositions immunogènes et/ou vaccinales selon l'invention destinées à la prévention et/ou au traitement d'infection par *Streptococcus* ou par un micro-organisme associé seront choisies parmi les compositions immunogènes et/ou vaccinales comprenant un polypeptide ou un de ses fragments correspondant à une protéine, ou un de ses fragments, de l'enveloppe cellulaire de *Streptococcus*. Les

compositions vaccinales comprenant des séquences nucléotidiques comprendront de préférence également des séquences nucléotidiques codant pour un polypeptide ou un de ses fragments correspondant à une protéine, ou un de ses fragments, de l'enveloppe cellulaire de *Streptococcus*.

5 Les polypeptides de l'invention ou leurs fragments entrant dans les compositions immunogènes selon l'invention peuvent être sélectionnés par des techniques connues de l'homme de l'art comme par exemple sur la capacité desdits polypeptides à stimuler les cellules T, qui se traduit par exemple par leur prolifération ou la sécrétion d'interleukines, ou qui aboutit à la production d'anticorps dirigés contre lesdits
10 polypeptides.

Chez la souris, chez laquelle une dose pondérale de la composition vaccinale comparable à la dose utilisée chez l'homme est administrée, la réaction anticorps est testée par prélèvement du sérum suivi d'une étude de la formation d'un complexe entre les anticorps présents dans le sérum et l'antigène de la composition vaccinale, selon les
15 techniques usuelles.

Selon l'invention, lesdites compositions vaccinales seront de préférence en association avec un véhicule pharmaceutiquement acceptable et, le cas échéant, avec un ou plusieurs adjuvants de l'immunité appropriés.

Aujourd'hui, divers types de vaccins sont disponibles pour protéger l'homme
20 contre des maladies infectieuses : micro-organismes vivants atténués (*M. bovis* - BCG pour la tuberculose), micro-organismes inactivés (virus de la grippe), des extraits acellulaires (*Bordetella pertussis* pour la coqueluche), protéines recombinées (antigène de surface du virus de l'hépatite B), des polysides (pneumocoques). Des vaccins préparés à partir de peptides de synthèse ou de micro-organismes génétiquement
25 modifiés exprimant des antigènes hétérologues sont en cours d'expérimentation. Plus récemment encore, des ADNs plasmidiques recombinés portant des gènes codant pour des antigènes protecteurs ont été proposés comme stratégie vaccinale alternative. Ce type de vaccination est réalisé avec un plasmide particulier dérivant d'un plasmide de *E. coli* qui ne se réplique pas *in vivo* et qui code uniquement pour la protéine vaccinante.
30 Des animaux ont été immunisés en injectant simplement l'ADN plasmidique nu dans le muscle. Cette technique conduit à l'expression de la protéine vaccinale *in situ* et à une réponse immunitaire de type cellulaire (CTL) et de type humoral (anticorps). Cette double induction de la réponse immunitaire est l'un des principaux avantages de la technique de vaccination avec de l'ADN nu.

Les compositions vaccinales comprenant des séquences nucléotidiques ou des vecteurs dans lesquels sont insérées lesdites séquences, sont notamment décrites dans la demande internationale N° WO 90/11092 et également dans la demande internationale N° WO 95/11307.

5 La séquence nucléotidique constitutive de la composition vaccinale selon l'invention peut être injectée à l'hôte après avoir été couplée à des composés qui favorisent la pénétration de ce polynucléotide à l'intérieur de la cellule ou son transport jusqu'au noyau cellulaire. Les conjugués résultants peuvent être encapsulés dans des microparticules polymères, comme décrit dans la demande internationale N° WO
10 94/27238 (Medisorb Technologies International).

Selon un autre mode de réalisation de la composition vaccinale selon l'invention, la séquence nucléotidique, de préférence un ADN, est complexée avec du DEAE-dextran, avec des protéines nucléaires, avec des lipides ou encapsulée dans des liposomes ou encore introduite sous la forme d'un gel facilitant sa transfection dans les
15 cellules. Le polynucléotide ou le vecteur selon l'invention peut aussi être en suspension dans une solution tampon ou être associé à des liposomes.

Avantageusement, un tel vaccin sera préparé conformément à la technique décrite par Tacson et al. ou Huygen et al. en 1996 ou encore conformément à la technique décrite par Davis et al. dans la demande internationale N° WO 95/11307.

20 Un tel vaccin peut être également préparé sous la forme d'une composition contenant un vecteur selon l'invention, placée sous le contrôle d'éléments de régulation permettant son expression chez l'homme ou l'animal. On pourra par exemple utiliser, en tant que vecteur d'expression *in vivo* de l'antigène polypeptidique d'intérêt, le plasmide pcDNA3 ou le plasmide pcDNA1/neo, tous les deux commercialisés par Invitrogen (R
25 & D Systems, Abingdon, Royaume-Uni). Un tel vaccin comprendra avantageusement, outre le vecteur recombinant, une solution saline, par exemple une solution de chlorure de sodium.

On entend désigner par véhicule pharmaceutiquement acceptable, un composé ou une combinaison de composés entrant dans une composition pharmaceutique ou
30 vaccinale ne provoquant pas de réactions secondaires et qui permet par exemple la facilitation de l'administration du composé actif, l'augmentation de sa durée de vie et/ou de son efficacité dans l'organisme, l'augmentation de sa solubilité en solution ou encore l'amélioration de sa conservation. Ces véhicules pharmaceutiquement acceptables sont

bien connus et seront adaptés par l'homme de l'art en fonction de la nature et du mode d'administration du composé actif choisi.

En ce qui concerne les formulations vaccinales, celles-ci peuvent comprendre des adjuvants de l'immunité appropriés qui sont connus de l'homme de l'art, comme par exemple l'hydroxyde d'aluminium, un représentant de la famille des muramyl peptides
5 comme un des dérivés peptidiques du N-acétyl-muramyl, un lysat bactérien, ou encore l'adjuvant incomplet de Freund.

De préférence, ces composés seront administrés par voie systémique, en particulier par voie intraveineuse, par voie intramusculaire, intradermique ou sous-
10 cutanée, ou par voie orale. De manière plus préférée, la composition vaccinale comprenant des polypeptides selon l'invention, sera administrée à plusieurs reprises, de manière étalée dans le temps, par voie intradermique ou sous-cutanée.

Leurs modes d'administration, posologies et formes galéniques optimaux peuvent être déterminés selon les critères généralement pris en compte dans
15 l'établissement d'un traitement adapté à un patient comme par exemple l'âge ou le poids corporel du patient, la gravité de son état général, la tolérance au traitement et les effets secondaires constatés.

Enfin, l'invention comprend l'utilisation d'une composition selon l'invention, pour le traitement ou la prévention de maladies induites ou aggravées par la présence de
20 *Streptococcus*.

Par ailleurs, la présente invention a également pour objet une banque d'ADN génomique d'une bactérie du genre *Streptococcus*, de manière préférée, *Streptococcus agalactiae*, de manière préférée la CIP 82.45 (ATCC 12403).

La banque d'ADN génomique décrite dans la présente invention, en particulier
25 la banque déposée à la CNCM le 28 décembre 2000 sous le numéro d'ordre N° I-2610, recouvre en effet *Streptococcus agalactiae* CIP 82.45 (ATCC 12403). Toutefois, si certaines régions n'ont pas pu être clonées dans ladite banque, en raison de problèmes de létalités chez *Escherichia coli*, ces régions peuvent facilement être amplifiées et identifiées par l'homme du métier, en utilisant des oligonucléotides spécifiques des
30 séquences des extrémités des différents clones qui forment les contigs.

La présente invention concerne également les méthodes pour l'isolement d'un polynucléotide d'intérêt présent chez une souche de *Streptococcus agalactiae* et absente chez une autre souche, qui utilise au moins une banque d'ADN basée par exemple sur un plasmide pSYX34 contenant un fragment du génome de *Streptococcus agalactiae*.

La méthode selon l'invention pour l'isolement d'un polynucléotide d'intérêt peut comprendre les étapes suivantes :

a) isoler au moins un polynucléotide contenu dans un clone de la banque d'ADN d'origine de *Streptococcus agalactiae*,

5 b) isoler :

- au moins un polynucléotide génomique ou ADNc d'une bactérie *Streptococcus*, ladite bactérie *Streptococcus* appartenant à une souche différente de la souche utilisée pour la construction de la banque d'ADN de l'étape a) ou, de façon alternative,

10 - au moins un polynucléotide contenu dans un clone d'une banque d'ADN préparé à partir du génome d'une *Streptococcus* qui est différente de la souche *Streptococcus agalactiae* utilisée pour la construction de la banque d'ADN de l'étape a) ;

c) hybrider le polynucléotide de l'étape a) au polynucléotide de l'étape b) ;

d) sélectionner les polynucléotides de l'étape a) qui n'ont pas formé de complexe
15 d'hybridation avec les polynucléotides de l'étape b) ;

e) caractériser le polynucléotide sélectionné.

On peut préparer le polynucléotide de l'étape a) par la digestion d'au moins un clone recombinant avec une enzyme de restriction appropriée, et de façon optionnelle, l'amplification de l'insert polynucléotide qui en résulte.

20 Ainsi, la méthode de l'invention permet à l'homme du métier d'effectuer des études génomiques comparatives entre les différentes souches ou espèces du genre *Streptococcus*, par exemple entre les souches pathogéniques et leurs équivalents non pathogènes.

En particulier, il est possible d'étudier et de déterminer les régions de
25 polymorphisme entre lesdites souches.

L'invention comprend aussi un procédé d'identification de séquence spécifique de *Streptococcus agalactiae*, caractérisé par l'alignement de séquences nucléotidiques de *Streptococcus agalactiae* selon l'invention et le traitement de données obtenues par cet alignement pour isoler les séquences spécifiques.

30

La présente invention concerne également l'utilisation des séquences nucléiques ou des polypeptides selon la présente invention :

- pour la sécrétion de protéines,
- comme facteurs de virulence,

- pour le contrôle par l'intermédiaire du quorum-sensing,
 - pour l'identification de cibles pour les maladies humaines dont *Streptococcus agalactiae* est un modèle, et
 - pour l'identification de cibles contre les bactéries Gram positives pathogènes
- 5 par la méthode de génomique soustractive (comme par exemple par comparaison avec des bactéries Gram positives non pathogènes).

EXEMPLES

Exemple 1 : Matériels et méthode

- 10 La stratégie de séquençage du génome de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) repose sur un séquençage alléatoire (shot-gun). La première étape de ce travail consiste à cloner l'ADN génomique de la bactérie *Streptococcus agalactiae* dans différents vecteurs (plasmides et BAC).

Matériels et méthodes.

- 15 1. Construction des banques :

a/ Banque de petits fragments dans le vecteur pcDNA2.1

- L'ADN chromosomique de la souche de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) a été préparé par une méthode classique incluant un traitement à la protéinase K et une extraction au phénol (9). Environ 10 µg d'ADN ont été cassés par
- 20 nébulisation (1 minute sous une pression de 1 bar) (4). Les extrémités des fragments d'ADN ont été rendues franches en faisant agir la DNA-polymérase du bactériophage T4 pendant 15 minutes à 37°C en présence des 4 nucléotides tri-phosphate. L'enzyme a été inactivée par une incubation de 15 mn à 75°C. Des adaptateurs (Invitrogen Cat. N° 408-18) ont ensuite été ligaturés à ces extrémités. Après ligation, les fragments
- 25 d'ADN chromosomiques ayant une taille entre 1 000 et 3 000 paires de bases ont été purifiés après électrophorèse sur gel d'agarose. Le vecteur utilisé pour la construction de la banque, pcDNA2.1 (Invitrogen), a été digéré par l'enzyme BstX1 et purifié par geneclon (BIO-101) après électrophorèse sur gel d'agarose. L'ADN chromosomique et le vecteur purifié ont été ligaturés par action de la ligase du bactériophage T4. Le
- 30 mélange de ligation a été introduit par transformation dans la souche d'*Escherichia coli* XL2-blue (Stratagene). Environ 4000 colonies sont obtenues par µl du mélange de ligation.

b/ Construction d'une banque de fragments de taille moyenne (banque déposée à la CNCM sous le numéro I-2610)

Afin de limiter la fréquence des événements de co-ligation et de clones sans insert, la méthode du remplissage partiel du site de coupure a été utilisée (7).

L'ADN chromosomique de la souche de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) a été digéré de manière partielle par l'enzyme de restriction Sau3A en utilisant une gamme de dilution de l'enzyme. La fraction présentant la gamme de tailles souhaitée (entre 2 et 12 kbases) a été précipitée et les extrémités ont été remplies partiellement par le fragment de Klenow de la DNA polymérase d'*E. coli* en présence de dTTP et dCTP. Après purification sur gel d'agarose les fragments d'ADN chromosomique ayant une taille comprise entre 3 et 6 kbases ont été ligués au vecteur pSYX34 (7) digéré par l'enzyme SalI remplie partiellement par l'enzyme de Klenow en présence de dATP et dGTP. Le mélange de ligation a été introduit par transformation dans la souche d'*Escherichia coli* XL10-kan (Stratagene) et étalement sur milieu LB contenant du chloramphénicol à une concentration de 20 mg/l. Environ 500 colonies sont obtenues par µl du mélange de ligation.

2. Préparation des plasmides et séquençage

Les plasmides ont été préparés par une méthode semi-automatique de préparation développée au laboratoire GMP basé sur la méthode de lyse alcaline (2). Les inserts chromosomiques ont été séquencés à partir de leurs deux extrémités en utilisant les primer T7 et universel en suivant les recommandations du fournisseur (PE-biosystems). Les séquences ont été déterminées en utilisant un séquenceur automatique de type 3700 (PE-Biosystem).

3. Assemblage des séquences

Les séquences ont été assemblées en utilisant l'ensemble de logiciel développé à l'Université de Washington, Phred, Phrap et Consed (5, 8). La finition de la séquence a été réalisée en utilisant l'ensemble de logiciel GMPTB (7). L'étape de finition correspond au reséquençage des régions où la séquence est peu sûre et le séquençage des régions situées entre les contigs. Elle a été réalisée en séquençant des produits de PCR correspondant à ces régions identifiées par un expert en utilisant les logiciels Consed (8) et GMPTB (7). Les séquences des oligonucléotides ont été définies en utilisant les logiciels Consed et Primo (8, 10).

4. Annotation des séquences

L'identification des phases codantes (CDS) a été réalisée en utilisant l'ensemble de logiciel GMPTB (L. Frangeul et al. non publié). Ce programme combine les résultats de différentes méthodes : (i) l'identification de phases ouvertes de lecture et

leur tri en fonction de leur taille, (ii) l'analyse de la probabilité d'être codante en utilisant le logiciel Genemark (11), (iii) l'identification d'un début de traduction (codon d'initiation et séquence de fixation du ribosome), (iv) similarité de la séquence protéique déduite avec les séquences protéiques contenues dans les banques de séquence en utilisant le logiciel BLASTP.

Les fonctions des protéines codées par les phases codantes identifiées indiquées au tableau 1 ont été prédites par l'analyse des résultats de recherche de similarités dans les banques en utilisant le logiciel BLASTP (1).

Exemple 2 : Description scientifique de la banque de BAC de *Streptococcus agalactiae*

10 CIP 82.45 (ATCC 12403) déposée à la CNCM le 28 décembre 2000 sous le numéro I-2610.

Collection de clones d'*Escherichia coli* DH 10B™ (Calvin et al., J. Bacteriol. 170, 2796, 1988) contenant des fragments d'ADN génomique de la bactérie *Streptococcus agalactiae* souche NEM 316, CIP 82.45 (ATCC 12403), clonés dans le vecteur pSYX34 (Xu et al., Biotechniques, 17:57, 1990). Le vecteur a été digéré Sall et partiellement rempli à l'aide de l'ADN polymérase afin de produire des extrémités 5'-TC. Des fragments génomiques aléatoires de 3-6 kb provenant de *Streptococcus agalactiae* (souche NEM 316, CIP 82-45, ATCC 12403) ont été digérés partiellement par Sau3A et partiellement remplis afin de produire des extrémités 5'-GA. Après
15
20 ligature *in vitro* et transformation, des clones résistant au chloromphénicol ont été sélectionnés. Environ 5000 clones ont été assemblés, suspendus dans 15 ml de milieu L et congelés.

Exemple 3 : Les protéines de surface de *Streptococcus agalactiae* NEM316

Les protéines de surface des bactéries pathogènes, et plus particulièrement les
25 protéines dites du type LPXTG (Navarre et Schneewind, Microbial. Mol. Biol. Rev. 63 174-229), jouent un rôle crucial lors du processus infectieux en permettant notamment des interactions entre le microorganisme et les cellules de l'hôte et/ou l'échappement au système immunitaire. Les inventeurs ont donc focalisé leur étude sur ce type de protéines qui présentent la particularité d'être liées covalamment au peptidoglycane via le motif
30 d'ancrage carboxylique LPXTG. Cette réaction est catalysée par une enzyme bifonctionnelle (endopeptidase-transpeptidase) appelée sortase. L'étude du rôle de ces protéines dans la virulence de *S. agalactiae* a été réalisée par 2 approches

complémentaires (construction d'un mutant sortase-déficient, inactivation des gènes codant pour des protéines de type LPXTG).

- Le gène *srtA* de *S. agalactiae* NEM316 (IPF N°1268).

Le mutant MEM1979, déposé le 24 avril 2002 à la CNCM sous le numéro I-
5 2861, est une souche mutante dérivée de NEM316 (CIP 82.45, ATCC 12403) dans laquelle l'IPF 1268 a été inactivé.

L'analyse du génome de NEM316 nous a permis de caractériser un gène *srtA* présentant homologie respectivement 55 % et 30 % d'identité avec les sortases de *Streptococcus gordonii* et de *Staphylococcus aureus*. Ce gène a été inactivé par
10 insertion-inactivation et nous avons montré que le mutant ainsi construit n'adhérait plus à des cellules épithéliales humaines pulmonaires (A549) et utérines (HeLa). Ces résultats suggèrent que les protéines du type LPXTG de *S. agalactiae* jouent un rôle dans la virulence de cette bactérie, en permettant notamment son adhésion avec les cellules de l'hôte.

15 - Recensement des protéines du type LPXTG de *S. agalactiae* NEM316.

Une analyse *in silico* du génome de NEM316 a révélé la présence de 30 protéines de surface putatives possédant le motif d'ancrage de type LPXTG (Tableau 6). Nous avons étudié par amplification PCR, en utilisant des amorces spécifiques, la distribution des gènes codant pour 21 protéines du type LPXTG dans une collection de
20 99 souches non-redondantes de *S. agalactiae* responsables d'infections non-invasives (70 souches provenant de portage ou d'infection urinaire) et invasives (29 souches provenant d'hémoculture ou de méningite). Cette étude a montré que 6 de ces gènes (IPN N° 1503, 678, 2192, 1861, 584, 280) étaient présents dans toutes les souches de notre échantillonnage (Tableau 7). Deux de ces six gènes (IPF N°678 et 1503) ont été
25 inactivés par insertion-inactivation et les mutants correspondants NEM2056 et NEM2057 présentaient une adhérence diminuée avec les cellules épithéliales A549 (Tableau 5).

Le mutant NEM2056, déposé le 24 avril 2002 à la CNCM sous le numéro I-
2862, est une souche mutante dérivée de NEM316 (CIP 82.45, ATCC 12403) dans
30 laquelle l'IPF 678 a été inactivé.

Le mutant NEM2057, déposé le 24 avril 2002 à la CNCM sous le numéro I-
2863, est une souche mutante dérivée de NEM316 (CIP 82.45, ATCC 12403) dans laquelle l'IPF 1503 a été inactivé.

- Cibles vaccinales.

La présence des gènes IPF N°1503, 678, 2192, 1861, 584, 280 chez toutes les souches de *S. agalactiae* testées font des protéines correspondantes des cibles vaccinales de choix pour le développement d'un vaccin anti-*S. agalactiae*.

5

TABLEAU 1: Liste des phases codantes annotées identifiées par l'analyse des séquences des contigs

10

SEQ ID No.	Prot No.	Localisation et sens sur contig	Résultats Blastp sur des banques de protéines non redondantes % Homologie /	Commentaires
SeqID 140	SA-1.2	Contig137 (42738-43340 p)	74	Identities = 129/202 (63%), Positives = 156/202 (76%), Gaps = 8/202 (3%) ref NP_072029.1 traG-related protein [Enterococcus faecalis] gb AAG40447.1 (AE002565) traG-related protein [Enterococcus faecalis] Length = 423
SeqID 141	SA-10.1	Contig137 (33454-33681 m)	No Hits found	
SeqID 142	SA-1000.1	Contig127 (13387-14454 p)	75	Identities = 220/354 (62%), Positives = 274/354 (77%), Gaps = 1/354 (0%) sp Q48677 PEPA_LALCLC GLUTAMYL-AMINOPEPTIDASE emb CAA56994.1 (X81089) glutamyl-aminopeptidase [Lactococcus lactis] Length = 355
SeqID 143	SA-1001.1	Contig127 (12912-13202 p)	No Hits found	
SeqID 144	SA-1002.1	Contig127 (12518-12802 m)	No Hits found	
SeqID 145	SA-1003.1	Contig127 (12198-12407 m)	65	Identities = 49/100 (49%), Positives = 71/100 (71%), Gaps = 4/100 (4%) dbj BAB06972.1 (AP001518) thioredoxin H1 [Bacillus halodurans] Length = 106
SeqID 146	SA-1004.1	Contig127 (11539-12165 m)	60	Identities = 79/196 (40%), Positives = 126/196 (63%), Gaps = 1/196 (0%) pir A69999 phenylalanyl-tRNA synthetase (beta subunit) homolog ytpR - Bacillus subtilis gb AAC00291.1 (AF008220) YtpR [Bacillus subtilis] emb CAB14960.1 (Z99119) similar to phenylalanyl-tRNA synthetase (beta subunit) [Bacillus subtilis] Length = 201
SeqID 147	SA-1006.1	Contig127 (10769-11485 p)	47	Identities = 69/243 (28%), Positives = 117/243 (47%), Gaps = 15/243 (6%) sp Q02148 YH16_LACLA HYPOTHETICAL 30.7 KDA PROTEIN IN HISD-HISB INTERGENIC REGION (ORF6) pir F45734 orf6 3 to hisD - Lactococcus lactis subsp. lactis gb AAB81904.1 (U92974) unknown [Lactococcus lactis] Length = 263

SeqID 148	SA-1007.1	Contig127 (10293-10688 m)	69	Identities = 65/141 (46%), Positives = 93/141 (65%), Gaps = 10/141 (7%) gb AAF74079.1 (AF212845) putative single stranded binding protein [Lactococcus lactis bacteriophage ul36] Length = 141
SeqID 149	SA-101.1	Contig131 (21108-21947 p)	82	Identities = 201/279 (72%), Positives = 231/279 (82%) gb AAC61484.1 (AF082738) ABC transporter ATP-binding protein [Streptococcus pyogenes] Length = 296
SeqID 150	SA-1010.1	Contig127 (9526-10170 m)	33	Identities = 63/179 (35%), Positives = 91/179 (50%), Gaps = 2/179 (1%) emb CAC13072.1 (AL445503) putative hydrolase [Streptomyces coelicolor A3(2)] Length = 238
SeqID 151	SA-1012.1	Contig127 (7754-9499 m)	60	Identities = 254/585 (43%), Positives = 371/585 (63%), Gaps = 9/585 (1%) pir A69655 two-component sensor histidine kinase lytS-involved - Bacillus subtilis emb CAA99610.1 (Z75208) autolysin sensor kinase [Bacillus subtilis] emb CAB14853.1 (Z99118) two-component sensor histidine kinase [Bacillus subtilis] Length = 593
SeqID 152	SA-1013.1	Contig127 (7033-7779 m)	62	Identities = 105/244 (43%), Positives = 157/244 (64%), Gaps = 6/244 (2%) pir B69655 two-component response regulator lytT-involved - Bacillus subtilis emb CAA99611.1 (Z75208) autolysin response regulator [Bacillus subtilis] emb CAB14852.1 (Z99118) two-component response regulator [Bacillus subtilis] Length = 241
SeqID 153	SA-1014.1	Contig127 (6936-7100 p)	No Hits found	
SeqID 154	SA-1016.1	Contig127 (6408-6863 m)	51	Identities = 50/110 (45%), Positives = 82/110 (74%), Gaps = 2/110 (1%) pir C69983 conserved hypothetical protein ysbA - Bacillus subtilis emb CAA99612.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14851.1 (Z99118) similar to hypothetical proteins from B. subtilis [Bacillus subtilis] Length = 146

SeqID 155	SA-1017.1	Contig127 (5678-6406 m)	64	Identities = 120/240 (50%), Positives = 159/240 (66%), Gaps = 10/240 (4%) pir D69983 conserved hypothetical protein ysbB - Bacillus subtilis emb CAA99613.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14850.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 231
SeqID 156	SA-1018.1	Contig127 (3807-5435 m)	90	Identities = 498/542 (91%), Positives = 518/542 (94%) gb AAC67217.1 (U78968) surface lipoprotein DppA [Streptococcus pyogenes] Length = 542
SeqID 157	SA-1019.1	Contig127 (2717-3694 m)	96	Identities = 302/325 (92%), Positives = 317/325 (96%) gb AAC67218.1 (U78968) transmembrane transport protein DppB [Streptococcus pyogenes] Length = 325
SeqID 158	SA-102.1	Contig131 (20566-21108 p)	69	Identities = 119/165 (72%), Positives = 134/165 (81%), Gaps = 5/165 (3%) gb AAC61483.1 (AF082738) phosphotidylglycerophosphate synthase [Streptococcus pyogenes] Length = 165
SeqID 159	SA-1020.1	Contig127 (1899-2720 m)	92	Identities = 255/274 (93%), Positives = 263/274 (95%), Gaps = 1/274 (0%) gb AAC67219.1 (U78968) transmembrane transport protein DppC [Streptococcus pyogenes] Length = 274
SeqID 160	SA-1021.1	Contig127 (1084-1887 m)	95	Identities = 254/267 (95%), Positives = 262/267 (97%) gb AAC67220.1 (U78968) ATPase protein DppD [Streptococcus pyogenes] Length = 267
SeqID 161	SA-1022.1	Contig127 (474-1100 m)	91	Identities = 185/205 (90%), Positives = 195/205 (94%) gb AAC67221.1 (U78968) ATPase protein DppE [Streptococcus pyogenes] Length = 208

SeqID 162	SA-1023.1	Contig127 (1-192 m)	62	<p>Identities = 35/58 (60%), Positives = 44/58 (75%) sp P36672 PTTB_ECOLI_PTS_SYSTEM, TREHALOSE-SPECIFIC IIBC COMPONENT (EIIBC-TRE) (TREHALOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EII-TRE) pir C65236 phosphotransferase system trehalose permease - Escherichia coli (strain K-12) gb AAC77197.1 (AE000495) PTS system enzyme II, trehalose specific [Escherichia coli K12] Length = 473</p>
SeqID 163	SA-1024.2	Contig113 (19147-19281 p)	No Hits found	
SeqID 164	SA-1025.2	Contig113 (18773-19144 p)	53	<p>Identities = 37/124 (29%), Positives = 67/124 (53%), Gaps = 2/124 (1%) gb AAC98903.1 (AF023181) DivIC homolog [Listeria monocytogenes] Length = 128</p>
SeqID 165	SA-1026.1	Contig113 (18514-18786 p)	75	<p>Identities = 49/84 (58%), Positives = 70/84 (83%) sp P37557 YABO_BACSU_HYPOTHETICAL_9.7_KDA_PROTEIN IN MFD-DIVIC INTERGENIC REGION pir S66089 conserved hypothetical protein yabO - Bacillus subtilis dbj BAA05294.1 (D26185) unknown [Bacillus subtilis] emb CAB11835.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 86</p>
SeqID 166	SA-1028.1	Contig113 (14726-18223 p)	62	<p>Identities = 523/1051 (49%), Positives = 733/1051 (68%), Gaps = 20/1051 (1%) gb AAD03810.1 (AF054624) transcription-repair coupling factor [Lactobacillus sakei] Length = 1045</p>
SeqID 167	SA-1029.1	Contig113 (14154-14729 p)	60	<p>Identities = 89/187 (47%), Positives = 127/187 (67%), Gaps = 2/187 (1%) dbj BAB03787.1 (AP001507) stage V sporulation protein C (peptidyl-tRNA hydrolase) [Bacillus halodurans] Length = 185</p>
SeqID 168	SA-103.1	Contig131 (19160-20443 p)	62	<p>Identities = 207/424 (48%), Positives = 277/424 (64%), Gaps = 3/424 (0%) gb AAC61481.1 (AF082738) unknown [Streptococcus pyogenes] Length = 429</p>

SeqID 169	SA-1030.1	Contig113 (12955-14070 p)	84	Identities = 261/371 (70%), Positives = 313/371 (84%), Gaps = 5/371 (1%) dbj BAB07770.1 (AP001520) GTP-binding protein [Bacillus halodurans] Length = 366
SeqID 170	SA-1031.1	Contig113 (12592-12795 p)	No Hits found	
SeqID 171	SA-1032.1	Contig113 (12193-12390 p)	82	Identities = 46/63 (73%), Positives = 57/63 (90%) gb AAC45338.1 (AF000658) ORFX [Streptococcus pneumoniae] Length = 64
SeqID 172	SA-1033.1	Contig113 (11290-12183 p)	43	Identities = 83/298 (27%), Positives = 138/298 (45%), Gaps = 35/298 (11%) pir H69995 hypothetical protein ytlR - Bacillus subtilis gb AAC00282.1 (AF008220) YtlR [Bacillus subtilis] emb CAB14972.1 (Z99119) ytlR [Bacillus subtilis] Length = 309
SeqID 173	SA-1034.1	Contig113 (10096-11232 p)	80	Identities = 278/378 (73%), Positives = 324/378 (85%) sp O06672 DP3B_STRPN DNA POLYMERASE III, BETA CHAIN gb AAC45337.1 (AF000658) beta subunit of DNA polymerase III [Streptococcus pneumoniae] Length = 378
SeqID 174	SA-1035.1	Contig113 (8580-9941 p)	82	Identities = 353/455 (77%), Positives = 401/455 (87%), Gaps = 6/455 (1%) gb AAF71535.1 AF255728_1 (AF255728) chromosomal initiator protein DnaA [Streptococcus pyogenes] Length = 451
SeqID 175	SA-1036.1	Contig113 (7608-8381 p)	71	Identities = 138/257 (53%), Positives = 188/257 (72%), Gaps = 5/257 (1%) gb AAC45335.1 (AF000658) SPSPoJ [Streptococcus pneumoniae] Length = 252
SeqID 176	SA-1037.1	Contig113 (6275-7510 p)	70	Identities = 222/396 (56%), Positives = 290/396 (73%), Gaps = 27/396 (6%) gb AAC45334.1 (AF000658) putative serine protease [Streptococcus pneumoniae] Length = 397
SeqID 177	SA-1038.1	Contig113 (5601-6080 m)	80	Identities = 111/159 (69%), Positives = 136/159 (84%) gb AAC44894.1 (U76218) unknown [Streptococcus pneumoniae] gb AAC45340.1 (AF000658) ORF1 [Streptococcus pneumoniae] Length = 159

SeqID 178	SA-1039.1	Contig113 (2443-5022 p)	21	Identities = 140/260 (53%), Positives = 182/260 (69%), Gaps = 6/260 (2%) gb AAC14608.1 (U95840) transmembrane protein Tmp5 [Lactococcus lactis] Length = 273
SeqID 179	SA-104.1	Contig131 (17914-19158 p)	73	Identities = 228/413 (55%), Positives = 307/413 (74%) gb AAC61480.1 (AF082738) unknown [Streptococcus pyogenes] Length = 414
SeqID 180	SA-1041.1	Contig113 (701-2320 p)	82	Identities = 354/542 (65%), Positives = 452/542 (83%), Gaps = 4/542 (0%) pir [E69861 ABC transporter (ATP-binding protein) homolog ykpA - Bacillus subtilis emb CAB13316.1 (Z99111) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] gb AAC24918.1 (AF012285) YkpA [Bacillus subtilis] Length = 540
SeqID 181	SA-1042.1	Contig113 (111-635 p)	56	Identities = 70/193 (36%), Positives = 119/193 (61%) pir [D70042 conserved hypothetical protein yvjA - Bacillus subtilis gb AAC67260.1 (AF017113) YvjA [Bacillus subtilis] emb CAB15545.1 (Z99122) similar to hypothetical proteins [Bacillus subtilis] Length = 281
SeqID 182	SA-1044.2	Contig130 (3069-3740 p)	62	Identities = 98/200 (49%), Positives = 139/200 (69%), Gaps = 1/200 (0%) gb AAF25544.1 AF109218_4 (AF109218) ThiE [Staphylococcus carnosus] Length = 212
SeqID 183	SA-1045.1	Contig130 (3867-5126 p)	70	Identities = 242/412 (58%), Positives = 303/412 (72%), Gaps = 2/412 (0%) gb AAF86297.1 (AF072894) UDP-N-acetylglucosamine-1-carboxyvinyl transferase [Listeria monocytogenes] Length = 439
SeqID 184	SA-1046.1	Contig130 (5210-5782 p)	44	Identities = 56/185 (30%), Positives = 86/185 (46%), Gaps = 4/185 (2%) dbj BAB04556.1 (AP001510) unknown conserved protein [Bacillus halodurans] Length = 197
SeqID 185	SA-1047.1	Contig130 (5763-7058 p)	62	Identities = 158/431 (36%), Positives = 270/431 (61%), Gaps = 6/431 (1%) dbj BAB06894.1 (AP001518) unknown conserved protein [Bacillus halodurans] Length = 435

SeqID 186	SA-1048.1	Contig130 (7081-7941 p)	44	Identities = 101/207 (48%), Positives = 128/207 (61%), Gaps = 31/207 (14%) gb AAC35914.1 (AF071085) methionine aminopeptidase A [Enterococcus faecalis] Length = 178
SeqID 187	SA-1049.1	Contig130 (7943-8863 p)	49	Identities = 83/276 (30%), Positives = 157/276 (56%), Gaps = 3/276 (1%) gb AAC35915.1 (AF071085) Orfde2 [Enterococcus faecalis] Length = 302
SeqID 188	SA-105.1	Contig131 (17322-17675 m)	61	Identities = 56/116 (48%), Positives = 74/116 (63%), Gaps = 3/116 (2%) pir JC4754 hypothetical 13.6k protein - Lactococcus lactis emb CAA61548.1 (X89367) orf121 [Lactococcus lactis] Length = 120
SeqID 189	SA-1050.1	Contig130 (8880-9314 m)	50	Identities = 52/161 (32%), Positives = 78/161 (48%), Gaps = 26/161 (16%) gb AAD54224.1 (AF143443_3 (AF143443) Mesh [Leuconostoc mesenteroides] Length = 160
SeqID 190	SA-1051.1	Contig130 (9517-10026 p)	47	Identities = 56/158 (35%), Positives = 91/158 (57%), Gaps = 8/158 (5%) emb CAA68244.1 (X99978) citrulline cluster-linked gene [Lactobacillus plantarum] Length = 168
SeqID 191	SA-1053.1	Contig130 (10166-12124 p)	70	Identities = 345/657 (52%), Positives = 462/657 (69%), Gaps = 8/657 (1%) sp O31498 DNLJ_BACSU DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE [NAD+]) pir F69794 DNA ligase (NAD+) (EC 6.5.1.2) - Bacillus subtilis emb CAB12482.1 (Z99107) similar to DNA ligase [Bacillus subtilis] Length = 668
SeqID 192	SA-1054.1	Contig130 (12136-13155 p)	54	Identities = 151/293 (51%), Positives = 193/293 (65%), Gaps = 5/293 (1%) pir F69795 conserved hypothetical protein yerQ - Bacillus subtilis emb CAB12492.1 (Z99107) similar to hypothetical proteins [Bacillus subtilis] Length = 303
SeqID 193	SA-1055.1	Contig130 (13159-15459 p)	47	Identities = 227/669 (33%), Positives = 363/669 (53%), Gaps = 66/669 (9%) prf 1808262A thermostable pullulanase [Bacillus stearothermophilus] Length = 658

SeqID 194	SA-1056.1	Contig130 (15665-17533 p)	60	Identities = 276/628 (43%), Positives = 377/628 (59%), Gaps = 20/628 (3%) sp P30537 GLGB_BACCL 1,4-ALPHA-GLUCAN BRANCHING ENZYME (GLYCOGEN BRANCHING ENZYME) pir B56639 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) - Bacillus caldolyticus emb CAA78440.1 (Z14057) 1,4-alpha-glucan branching enzyme [Bacillus caldolyticus] Length = 666
SeqID 195	SA-1057.1	Contig130 (17575-18714 p)	65	Identities = 196/352 (55%), Positives = 259/352 (72%) sp O08326 GLGC_BACST GLUCOSE-1-PHOSPHATE ADENYLTRANSFERASE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE) dbj BAA19589.1 (D87026) subunit of ADP-glucose pyrophosphorylase [Bacillus stearothermophilus] Length = 387
SeqID 196	SA-1058.1	Contig130 (18704-19837 p)	47	Identities = 105/353 (29%), Positives = 180/353 (50%), Gaps = 9/353 (2%) dbj BAB04805.1 (AP001510) required for glycogen biosynthesis [Bacillus halodurans] Length = 368
SeqID 197	SA-1059.2	Contig130 (19834-21264 p)	65	Identities = 221/475 (46%), Positives = 313/475 (65%), Gaps = 1/475 (0%) sp O08328 GLGA_BACST GLYCOGEN SYNTHASE (STARCH [BACTERIAL GLYCOGEN] SYNTHASE) dbj BAA19591.1 (D87026) bacterial glycogen synthase [Bacillus stearothermophilus] Length = 485
SeqID 198	SA-106.1	Contig131 (16210-17319 m)	78	Identities = 249/364 (68%), Positives = 300/364 (82%), Gaps = 1/364 (0%) sp P49999 REFC_STRPY RECF PROTEIN pir JC4077 recF protein - Streptococcus pyogenes gb AAA85783.1 (U07342) RecF protein [Streptococcus pyogenes] Length = 368
SeqID 199	SA-1060.1	Contig109 (28847-29194 m)	64	Identities = 71/87 (81%), Positives = 75/87 (85%) emb CAA13587.1 (AJ233894) xanthine phosphoribosyltransferase [Streptococcus pneumoniae] Length = 162
SeqID 200	SA-1061.1	Contig109 (28189-28557 p)	No Hits found	

SeqID 201	SA-1062.1	Contig109 (27573-28847 m)	68	Identities = 213/412 (51%), Positives = 292/412 (70%), Gaps = 5/412 (1%) sp P42086 PBUX_BACSU XANTHINE PERMEASE pir S51310 xanthine transport protein pbuX - Bacillus subtilis emb CAA58759.1 (X83878) xanthine permease [Bacillus subtilis] gb AAA96612.1 (L77246) transport protein [Bacillus subtilis] emb CAB14123.1 (Z99115) xanthine permease [Bacillus subtilis] Length = 438
SeqID 202	SA-1063.1	Contig109 (26686-27474 p)	44	Identities = 82/253 (32%), Positives = 119/253 (46%), Gaps = 10/253 (3%) sp P39608 YWCJ_BACSU HYPOTHETICAL 28.4 KD PROTEIN IN SACT-SACP INTERGENIC REGION pir S39703 nitrite transport protein homolog ywcJ - Bacillus subtilis emb CAA51604.1 (X73124) ipa-48r [Bacillus subtilis] emb CAB15832.1 (Z99123) alternate gene name: ipa-48r~similar to nitrite transporter [Bacillus subtilis] Length = 256
SeqID 203	SA-1064.1	Contig109 (25352-26587 p)	26	Identities = 80/162 (49%), Positives = 112/162 (68%) gb AAG18632.1 (AY007504) unknown [Streptococcus mitis] Length = 173
SeqID 204	SA-1065.1	Contig109 (24726-25328 p)	54	Identities = 93/160 (58%), Positives = 120/160 (74%), Gaps = 1/160 (0%) gb AAG18632.1 (AY007504) unknown [Streptococcus mitis] Length = 173
SeqID 205	SA-1066.1	Contig109 (23755-24708 p)	50	Identities = 103/317 (32%), Positives = 165/317 (51%), Gaps = 15/317 (4%) sp O83774 APBE_TREPA THIAMINE BIOSYNTHESIS LIPOPROTEIN APBE PRECURSOR pir C71281 conserved hypothetical protein TP0796 - syphilis spirochete gb AAC65759.1 (AE001250) conserved hypothetical protein [Treponema pallidum] Length = 362

SeqID 206	SA-1067.1	Contig109 (23439-23621 p)	55	<p>Identities = 27/60 (45%), Positives = 36/60 (60%) sp Q01468 XYLH_PSEPU 4-OXALOCROTTONATE TAUTOMERASE (4-OT) pir A43397 4-oxalocrotonate tautomerase (EC 5.3.2.-) xylH - Pseudomonas putida plasmid TOL pWWO gb AAA25694.1 (M94186) 4- oxalocrotonate tautomerase [Pseudomonas putida] gb AAA26046.1 (M95650) 4-oxalocrotonate tautomerase [Plasmid pWWO] prf 1916401D 4-oxalocrotonate tautomerase [Pseudomonas putida] Length = 63</p>
SeqID 207	SA-1068.1	Contig109 (22732-23301 m)	89	<p>Identities = 157/189 (83%), Positives = 175/189 (92%) sp P47848 KITH_STRGC THYMIDINE KINASE gb AAB02289.1 (L40415) thymidine kinase [Streptococcus gordonii] Length = 191</p>
SeqID 208	SA-1069.1	Contig109 (21618-22697 m)	76	<p>Identities = 211/351 (60%), Positives = 281/351 (79%), Gaps = 1/351 (0%) sp P45872 RF1_BACSU PEPTIDE CHAIN RELEASE FACTOR 1 (RF-1) pir S55437 translation releasing factor RF-1 - Bacillus subtilis emb CAA89884.1 (Z49782) peptide chain release factor 1 [Bacillus subtilis] emb CAB15718.1 (Z99122) peptide chain release factor 1 [Bacillus subtilis] Length = 356</p>
SeqID 209	SA-107.1	Contig131 (15332-16198 p)	78	<p>Identities = 195/277 (70%), Positives = 236/277 (84%) gb AAA56773.1 (U17382) putative multiple membrane domain protein; possible TTG initiation codon at position 1064, near putative RBS at position 1052 [Streptococcus pyogenes] Length = 277</p>
SeqID 210	SA-1071.1	Contig109 (20788-21618 m)	47	<p>Identities = 94/236 (39%), Positives = 132/236 (55%), Gaps = 12/236 (5%) dbj BAB07493.1 (AP001519) protoporphyrinogen oxidase [Bacillus halodurans] Length = 289</p>
SeqID 211	SA-1072.1	Contig109 (20199-20795 m)	57	<p>Identities = 71/199 (35%), Positives = 115/199 (57%), Gaps = 4/199 (2%) pir E72324 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD35934.1 AE001752_1 (AE001752) conserved hypothetical protein [Thermotoga maritima] Length = 335</p>

SeqID 212	SA-1073.1	Contig109 (18851-20107 m)	73	Identities = 242/417 (58%), Positives = 308/417 (73%), Gaps = 7/417 (1%) sp Q9WZH9 GLYA_THEMA SERINE HYDROXYMETHYLTRANSFERASE (SERINE METHYLASE) (SHMT) pir F72341 glycine hydroxymethyltransferase (EC 2.1.2.1) - Thermotoga maritima (strain MSB8) gb AAD35802.1 AE001743_4 (AE001743) serine hydroxymethyltransferase [Thermotoga maritima] Length = 427
SeqID 213	SA-1074.1	Contig109 (17869-18846 m)	15	Identities = 32/95 (33%), Positives = 53/95 (55%), Gaps = 1/95 (1%) emb CAB71302.1 (AJ130879) hypothetical protein [Clostridium sticklandii] Length = 95
SeqID 214	SA-1075.1	Contig109 (17265-17867 m)	36	Identities = 53/140 (37%), Positives = 78/140 (54%), Gaps = 2/140 (1%) pir F69900 hypothetical protein yocA [imported] - Bacillus subtilis gb AAB84433.1 (AF027868) RAS-related protein [Bacillus subtilis] emb CAB13805.1 (Z99114) similar to transposon-related protein [Bacillus subtilis] Length = 225
SeqID 215	SA-1076.1	Contig109 (15532-17253 m)	55	Identities = 196/570 (34%), Positives = 327/570 (56%), Gaps = 5/570 (0%) pir E72396 ABC transporter, ATP-binding protein - Thermotoga maritima (strain MSB8) gb AAD35375.1 AE001710_10 (AE001710) ABC transporter, ATP-binding protein [Thermotoga maritima] Length = 577
SeqID 216	SA-1077.1	Contig109 (13798-15531 m)	58	Identities = 219/552 (39%), Positives = 337/552 (60%), Gaps = 3/552 (0%) sp Q9WYC4 Y288_THEMA HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN TM0288 pir F72396 ABC transporter, ATP-binding protein - Thermotoga maritima (strain MSB8) gb AAD35376.1 AE001710_11 (AE001710) ABC transporter, ATP-binding protein [Thermotoga maritima] Length = 598
SeqID 217	SA-1078.2	Contig109 (13587-13841 m)	No Hits found	

SeqID 218	SA-1079.1	Contig109 (11805-13523 p)	91	Identities = 486/573 (84%), Positives = 527/573 (91%), Gaps = 1/573 (0%) emb CAB96418.1 (AJ243290) phosphoglucosyltransferase [Streptococcus thermophilus] Length = 572
SeqID 219	SA-108.1	Contig131 (14614-15276 p)	45	Identities = 51/215 (23%), Positives = 102/215 (46%), Gaps = 9/215 (4%) gb AAD04237.1 (AF007761) MutR [Streptococcus mutans] Length = 287
SeqID 220	SA-1080.1	Contig109 (11246-11695 p)	43	Identities = 49/190 (25%), Positives = 84/190 (43%), Gaps = 33/190 (17%) emb CAA80247.1 (Z22520) membrane protein [Bacillus acidopullulicus] Length = 183
SeqID 221	SA-1081.1	Contig109 (10535-11077 p)	66	Identities = 101/145 (69%), Positives = 122/145 (83%) sp Q54433 DFP_STRMU DNA/PANTOTHENATE METABOLISM FLAVOPROTEIN HOMOLOG gb AAC44502.1 (U48885) DNA/pantothenate metabolism flavoprotein [Streptococcus mutans] Length = 145
SeqID 222	SA-1082.1	Contig109 (9841-10542 p)	46	Identities = 74/232 (31%), Positives = 122/232 (51%), Gaps = 36/232 (15%) sp Q58323 DFP_METJA DNA/PANTOTHENATE METABOLISM FLAVOPROTEIN HOMOLOG pir JA64414 pantothenate metabolism flavoprotein MJ0913 - Methanococcus jannaschii gb AAB98918.1 (U67535) pantothenate metabolism flavoprotein (dfp) [Methanococcus jannaschii] Length = 403
SeqID 223	SA-1083.1	Contig109 (8866-9702 m)	43	Identities = 94/221 (42%), Positives = 134/221 (60%), Gaps = 2/221 (0%) db BAB07541.1 (AP001520) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 278
SeqID 224	SA-1084.1	Contig109 (7674-8873 m)	54	Identities = 162/311 (52%), Positives = 218/311 (70%), Gaps = 1/311 (0%) pir S49950 probable trimethylamine dehydrogenase (EC 1.5.99.7) - Mycoplasma capricolum (SGC3) (fragment) pir S77803 hypothetical protein MC012 - Mycoplasma capricolum emb CAA83700.1 (Z33015) similar to trimethylamine DH [Mycoplasma capricolum] Length = 311

SeqID 225	SA-1085.1	Contig109 (6691-7677 m)	57	Identities = 121/333 (36%), Positives = 190/333 (56%), Gaps = 12/333 (3%) dbj BAB04594.1 (AP001510) unknown conserved protein [Bacillus halodurans] Length = 335
SeqID 226	SA-1086.1	Contig109 (6330-6662 m)	41	Identities = 30/91 (32%), Positives = 53/91 (57%), Gaps = 1/91 (1%) gb AAF81345.1 AC007767_25 (AC007767) identical to a glycine cleavage system H-protein precursor from Arabidopsis thaliana gb P25855. It contains a glycine cleavage H-protein domain PF01597. ESTs gb R90208, gb AI994794, gb AA605324, gb IN38240, gb AV533336, gb AV534187, gb AA5> Length = 166
SeqID 227	SA-1087.2	Contig109 (5519-6337 m)	35	Identities = 71/171 (41%), Positives = 97/171 (56%), Gaps = 13/171 (7%) ref NP_054786.1 LRP16 protein [Homo sapiens] gb AAF15294.1 AF202922_1 (AF202922) LRP16 [Homo sapiens] gb AAH00270.1 AAH00270 (BC000270) LRP16 protein [Homo sapiens] gb AAH03188.1 AAH03188 (BC003188) LRP16 protein [Homo sapiens] Length = 243
SeqID 228	SA-1088.2	Contig138 (9701-11002 m)	56	Identities = 160/415 (38%), Positives = 251/415 (59%), Gaps = 13/415 (3%) sp Q45400 PTCC_BACST PTS SYSTEM, CELLOBIOSE-SPECIFIC IIC COMPONENT (EIIIC-CEL) (CELLOBIOSE-PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT) pir C49898 cellobiose phosphotransferase system celB - Bacillus stearothermophilus gb AAA17390.1 (U07818) cellobiose phosphotransferase enzyme II [Bacillus stearothermophilus] Length = 451
SeqID 229	SA-1089.1	Contig138 (11004-11324 m)	55	Identities = 46/100 (46%), Positives = 62/100 (62%), pir D82219 PTS system, cellobiose-specific IIB component VC1281 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF94440.1 (AE004207) PTS system, cellobiose-specific IIB component [Vibrio cholerae] Length = 101

SeqID 230	SA-109.1	Contig131 (13055-14539 p)	93	Identities = 461/493 (93%), Positives = 478/493 (96%) sp P50099 IMDH_STRPY INOSINE-5 -MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD) pir JC4372 IMP dehydrogenase (EC 1.1.1.205) - Streptococcus pyogenes gb AAB03846.1 (U26056) inosine monophosphate dehydrogenase [Streptococcus pyogenes] Length = 493
SeqID 231	SA-1090.1	Contig138 (11341-11670 m)	60	Identities = 42/102 (41%), Positives = 70/102 (68%) gb AAA23551.1 (M93570) PTS enzyme III cel [Escherichia coli] Length = 116
SeqID 232	SA-1091.2	Contig138 (11868-12857 m)	53	Identities = 98/309 (31%), Positives = 178/309 (56%), Gaps = 1/309 (0%) dbj BAB04499.1 (AP001509) transcriptional regulator [Bacillus halodurans] Length = 316
SeqID 233	SA-1092.2	Contig138 (12970-13725 m)	56	Identities = 96/243 (39%), Positives = 148/243 (60%), Gaps = 1/243 (0%) sp P76034 YCIT_ECOLI HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN OSMB-RNB INTERGENIC REGION pir G64876 hypothetical protein b1284 - Escherichia coli gb AAC74366.1 (AE000226) putative DEOR-type transcriptional regulator [Escherichia coli K12] Length = 249
SeqID 234	SA-1093.1	Contig138 (13846-14622 p)	54	Identities = 106/289 (36%), Positives = 146/289 (49%), Gaps = 50/289 (17%) sp P75794 PFLE_ECOLI PUTATIVE PYRUVATE FORMATE-LYASE 3 ACTIVATING ENZYME pir H64819 probable pyruvate formate-lyase activating enzyme (EC 1.97.1.4) 3 - Escherichia coli gb AAC73911.1 (AE000184) putative pyruvate formate-lyase 2 activating enzyme [Escherichia coli K12] dbj BAA35512.1 (D90720) Iron-dependent pyruvate formate-lyase-activating enzyme [Escherichia coli] Length = 308

SeqID 235	SA-1094.1	Contig138 (14669-15025 m)	56	Identities = 46/120 (38%), Positives = 71/120 (58%), Gaps = 11/120 (9%) dbj BAB07066.1 (AP001518) polyribonucleotide nucleotidyltransferase (general stress protein 13) [Bacillus halodurans] Length = 138
SeqID 236	SA-1095.1	Contig138 (15027-16385 m)	23	Identities = 81/174 (46%), Positives = 109/174 (62%), Gaps = 30/174 (17%) sp P87051 YDJ3_SCHPO PROBABLE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE C57A10.03 pir T38930 peptidylprolyl isomerase (EC 5.2.1.8) SPAC57A10.03 [similarity] - fission yeast (Schizosaccharomyces pombe) emb CAB08166.1 (Z94864) putative peptidyl-prolyl cis-trans isomerase [Schizosaccharomyces pombe] Length = 155
SeqID 237	SA-1096.1	Contig138 (16469-17122 m)	83	Identities = 155/209 (74%), Positives = 184/209 (87%) emb CAB54571.1 (AJ006393) response regulator [Streptococcus pneumoniae] Length = 210
SeqID 238	SA-1097.1	Contig138 (17103-18122 m)	70	Identities = 159/334 (47%), Positives = 239/334 (70%), Gaps = 5/334 (1%) emb CAB54570.1 (AJ006393) histidine kinase [Streptococcus pneumoniae] Length = 331
SeqID 239	SA-1098.1	Contig138 (18119-18814 m)	43	Identities = 53/230 (23%), Positives = 104/230 (45%), Gaps = 14/230 (6%) dbj BAB03323.1 (AB035448) hypothetical protein [Staphylococcus aureus] Length = 233
SeqID 240	SA-1099.1	Contig138 (18989-19231 p)	No Hits found	
SeqID 241	SA-11.1	Contig137 (32951-36100 p)	27	Identities = 170/605 (28%), Positives = 289/605 (47%), Gaps = 105/605 (17%) ref NP_053171.1 pXO2-16 [Bacillus anthracis] gb AAAF13621.1 AF188935_19 (AF188935) pXO2-16 [Bacillus anthracis] Length = 611
SeqID 242	SA-110.1	Contig131 (12416-12901 p)	56	Identities = 63/144 (43%), Positives = 93/144 (63%), Gaps = 4/144 (2%) emb CAA09426.1 (AJ010954) arginine repressor [Bacillus stearothermophilus] Length = 149

SeqID 243	SA-1100.1	Contig138 (18973-20928 m)	43	Identities = 219/380 (57%), Positives = 284/380 (74%), Gaps = 8/380 (2%) emb CAA10713.1 (AJ132604) hypothetical protein [Lactococcus lactis] Length = 378
SeqID 244	SA-1101.1	Contig138 (20928-21665 m)	69	Identities = 131/245 (53%), Positives = 177/245 (71%), Gaps = 4/245 (1%) emb CAA10712.1 (AJ132604) pppL protein [Lactococcus lactis] Length = 258
SeqID 245	SA-1102.1	Contig138 (21703-23025 m)	68	Identities = 222/434 (51%), Positives = 305/434 (70%), Gaps = 15/434 (3%) emb CAA10711.1 (AJ132604) sunL protein [Lactococcus lactis] Length = 424
SeqID 246	SA-1103.1	Contig138 (23015-23950 m)	67	Identities = 158/314 (50%), Positives = 215/314 (68%), Gaps = 5/314 (1%) dbj BAB06227.1 (AP001515) methionyl-tRNA formyltransferase [Bacillus halodurans] Length = 317
SeqID 247	SA-1104.1	Contig138 (23997-26387 m)	66	Identities = 377/807 (46%), Positives = 529/807 (64%), Gaps = 21/807 (2%) sp P94461 PRIA_BACSU PRIMOSOMAL PROTEIN N (REPLICATION FACTOR Y) pir A69682 primosomal replication factor Y priA - Bacillus subtilis emb CAA74261.1 (Y13937) putative PriA protein [Bacillus subtilis] emb CAB13444.1 (Z99112) primosomal replication factor Y (primosomal protein N%) [Bacillus subtilis] Length = 805
SeqID 248	SA-1105.2	Contig138 (26461-26775 m)	35	Identities = 25/57 (43%), Positives = 40/57 (69%), Gaps = 1/57 (1%) pir C69878 hypothetical protein yloH - Bacillus subtilis emb CAA74272.1 (Y13937) putative rpoZ protein [Bacillus subtilis] emb CAB13442.1 (Z99112) yloH [Bacillus subtilis] Length = 67
SeqID 249	SA-1107.1	Contig105 (16384-17661 p)	75	Identities = 260/416 (62%), Positives = 324/416 (77%) dbj BAB06905.1 (AP001518) argininosuccinate lyase [Bacillus halodurans] Length = 458

SeqID 250	SA-1108.1	Contig105 (15175-16365 p)	79	Identities = 262/396 (66%), Positives = 321/396 (80%), Gaps = 1/396 (0%) dbj BAB06906.1 (AP001518) argininosuccinate synthase (citrulline-aspartate ligase) [Bacillus halodurans] Length = 409
SeqID 251	SA-1109.1	Contig105 (14002-15021 p)	46	Identities = 97/307 (31%), Positives = 170/307 (54%), Gaps = 16/307 (5%) gb AAC62214.1 (AF049873) sensor protein [Lactococcus lactis] Length = 464
SeqID 252	SA-111.1	Contig131 (11879-12406 p)	60	Identities = 102/174 (58%), Positives = 141/174 (80%) pir JH0364 hypothetical protein 176 (SAGP 5 region) - Streptococcus pyogenes Length = 176
SeqID 253	SA-1110.1	Contig105 (13287-13958 p)	58	Identities = 86/231 (37%), Positives = 133/231 (57%), Gaps = 10/231 (4%) emb CAB64972.1 (AJ012050) VicR protein [Enterococcus faecalis] Length = 283
SeqID 254	SA-1111.1	Contig105 (12570-13256 p)	62	Identities = 87/225 (38%), Positives = 146/225 (64%), Gaps = 1/225 (0%) emb CAA05977.1 (AJ003195) ATP-binding subunit [Anabaena variabilis] Length = 244
SeqID 255	SA-1112.1	Contig105 (11482-12570 p)	50	Identities = 103/352 (29%), Positives = 191/352 (54%), Gaps = 9/352 (2%) pir F75581 hypothetical protein - Deinococcus radiodurans (strain R1) gb AAF12525.1 AE001863_150 (AE001863) hypothetical protein [Deinococcus radiodurans] Length = 353
SeqID 256	SA-1113.1	Contig105 (10314-11300 m)	61	Identities = 141/327 (43%), Positives = 204/327 (62%), Gaps = 4/327 (1%) sp P36944 IBSR_BACSU RIBOSE OPERON REPRESSOR pir E69690 transcription repressor of ribose operon - Bacillus subtilis emb CAB07467.1 (Z92953) repressor [Bacillus subtilis] emb CAB15608.1 (Z99122) transcriptional regulator (LacI family) [Bacillus subtilis] Length = 326

SeqID 257	SA-1114.1	Contig105 (9410-10321 m)	57	<p>Identities = 132/293 (45%), Positives = 176/293 (60%), Gaps = 4/293 (1%) sp P36945 RBSK_BACSU RIBOKINASE pir D69690 ribokinase (EC 2.7.1.15) - Bacillus subtilis emb CAB07465.1 (Z92953) ribokinase [Bacillus subtilis] emb CAB15609.1 (Z99122) ribokinase [Bacillus subtilis] Length = 293</p>
SeqID 258	SA-1115.1	Contig105 (9037-9435 m)	60	<p>Identities = 74/131 (56%), Positives = 96/131 (72%), Gaps = 1/131 (0%) sp P36946 RBSB_BACSU HIGH AFFINITY RIBOSE TRANSPORT PROTEIN RBSB pir J40464 ribose ABC transporter (membrane protein) rbsD - Bacillus subtilis emb CAA81050.1 (Z25798) Membrane Ribose Binding Protein [Bacillus subtilis] emb CAB07464.1 (Z92953) membrane transport protein [Bacillus subtilis] emb CAB15610.1 (Z99122) ribose ABC transporter (membrane protein) [Bacillus subtilis] Length = 131</p>
SeqID 259	SA-1116.1	Contig105 (7543-9021 m)	76	<p>Identities = 298/493 (60%), Positives = 377/493 (76%), Gaps = 1/493 (0%) sp P36947 RBSA_BACSU RIBOSE TRANSPORT ATP-BINDING PROTEIN RBSA pir H69689 ribose ABC transporter (ATP-binding protein) rbsA - Bacillus subtilis emb CAB07463.1 (Z92953) ATP-binding transport protein [Bacillus subtilis] emb CAB15611.1 (Z99122) ribose ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 493</p>
SeqID 260	SA-1118.1	Contig105 (6600-7541 m)	78	<p>Identities = 198/305 (64%), Positives = 252/305 (81%), Gaps = 1/305 (0%) pir B69690 ribose ABC transporter (permease) rbsC - Bacillus subtilis emb CAB07462.1 (Z92953) membrane transport protein [Bacillus subtilis] emb CAB15612.1 (Z99122) ribose ABC transporter (permease) [Bacillus subtilis] Length = 322</p>

SeqID 261	SA-1119.1	Contig105 (5579-6547 m)	59	<p>Identities = 143/301 (47%), Positives = 204/301 (67%), Gaps = 1/301 (0%) sp P36949 RBSB_BACSU D-RIBOSE-BINDING PROTEIN PRECURSOR pir A69690 ribose ABC transporter (ribose-binding protein) rbsB - Bacillus subtilis emb CAB07461.1 (Z92953) periplasmic substrate-binding protein [Bacillus subtilis] emb CAB15613.1 (Z99122) ribose ABC transporter (ribose-binding protein) [Bacillus subtilis]</p> <p>Length = 305</p>
SeqID 262	SA-112.1	Contig131 (10814-11482 m)	31	<p>Identities = 42/132 (31%), Positives = 74/132 (55%), Gaps = 3/132 (2%) pir H83123 hypothetical protein PA4181 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG07568.1 AE004834_8 (AE004834) hypothetical protein [Pseudomonas aeruginosa]</p> <p>Length = 239</p>
SeqID 263	SA-1120.1	Contig105 (4085-5539 p)	69	<p>Identities = 273/491 (55%), Positives = 353/491 (71%), Gaps = 19/491 (3%) sp P22250 SYE_BACSU GLUTAMYL-TRNA SYNTHETASE (GLUTAMATE--TRNA LIGASE) (GLURS) pir SYBSET glutamate--tRNA ligase (EC 6.1.1.17) gltX - Bacillus subtilis gb AAA22495.1 (M55073) transfer RNA-Gln synthetase [Bacillus subtilis] gb AAA21796.1 (L14580) glutamyl-tRNA synthetase [Bacillus subtilis] dbj BAA05326.1 (D26185) glutamyl-tRNA synthetase [Bacillus subtilis] emb CAB11868.1 (Z99104) tRNA synthetase [Bacillus subtilis] gb AAC31971.1 (U49789) glutamyl-tRNA synthetase [Bacillus subtilis]</p> <p>Length = 483</p>
SeqID 264	SA-1121.1	Contig105 (2557-3876 p)	65	<p>Identities = 201/443 (45%), Positives = 290/443 (65%) sp P77212 YKGC_ECOLI PROBABLE PYRIDINE NUCLEOTIDE-DISULFIDE OXIDOREDUCTASE IN EAEH-BETA INTERGENIC REGION pir H64756 probable mercury(II) reductase (EC 1.16.1.1) - Escherichia coli gb AAB18031.1 (U73857) similar to S. aureus mercury(II) reductase [Escherichia coli] gb AAC73407.1 (AE000137) putative oxidoreductase [Escherichia coli K12]</p> <p>Length = 450</p>

SeqID 265	SA-1122.1	Contig105 (1906-2403 p)	56	Identities = 69/162 (42%), Positives = 100/162 (61%), Gaps = 2/162 (1%) sp Q10612 YC84_MYCTU HYPOTHETICAL 18.2 KDA PROTEIN RV1284 pir H70771 hypothetical protein RV1284 - Mycobacterium tuberculosis (strain H37RV) emb CAA97750.1 (Z73419) hypothetical protein RV1284 [Mycobacterium tuberculosis] Length = 163
SeqID 266	SA-1123.1	Contig105 (406-1770 p)	78	Identities = 284/453 (62%), Positives = 358/453 (78%), Gaps = 4/453 (0%) sp P37572 RADA_BACSU DNA REPAIR PROTEIN RADA HOMOLOG (DNA REPAIR PROTEIN SMS HOMOLOG) pir S66116 DNA repair protein sms - Bacillus subtilis dbj BAA05321.1 (D26185) unknown [Bacillus subtilis] emb CAB11863.1 (Z99104) DNA repair protein homolog [Bacillus subtilis] Length = 458
SeqID 267	SA-1124.1	Contig105 (2-244 p)	64	Identities = 41/80 (51%), Positives = 52/80 (64%), Gaps = 5/80 (6%) gb AAAF74088.1 (AF212845) putative dUTPase [Lactococcus lactis bacteriophage ul36] Length = 139
SeqID 268	SA-1125.2	Contig120 (17546-17980 p)	44	Identities = 46/128 (35%), Positives = 70/128 (53%), Gaps = 5/128 (3%) gb AAK05804.1 AE006401_5 (AE006401) mannose-specific PTS system component IIAB (EC 2.7.1.69) [Lactococcus lactis subsp. lactis] Length = 329
SeqID 269	SA-1126.2	Contig120 (17983-19179 p)	54	Identities = 156/370 (42%), Positives = 219/370 (59%), Gaps = 3/370 (0%) dbj BAB05773.1 (AP001514) unsaturated glucuronyl hydrolase [Bacillus halodurans] Length = 375
SeqID 270	SA-1127.1	Contig120 (19234-19728 p)	52	Identities = 63/125 (50%), Positives = 89/125 (70%), Gaps = 1/125 (0%) gb AAC44679.1 (U65015) PTS permease for mannose subunit IIIMan C terminal domain [Vibrio furnissii] Length = 157
SeqID 271	SA-1129.1	Contig120 (19764-20630 p)	45	Identities = 93/295 (31%), Positives = 140/295 (46%), Gaps = 48/295 (16%) gb AAAF81084.1 AF228498_4 (AF228498) AgaW [Escherichia coli] Length = 259

SeqID 272	SA-113.1	Contig131 (9315-10547 m)	95	Identities = 375/411 (91%), Positives = 395/411 (95%), Gaps = 1/411 (0%) sp P16962 ARCA_STRPY ARGININE DEIMINASE (ADI) (ARGININE DIHYDROLASE) (AD) (STREPTOCOCCAL ACID GLYCOPROTEIN) pir A38835 streptococcal antitumor protein - Streptococcus pyogenes emb CAA39192.1 (X55659) antitumor protein [Streptococcus pyogenes] dbj BAA02938.1 (D13790) acid glycoprotein [Streptococcus pyogenes] Length = 411
SeqID 273	SA-1131.1	Contig120 (20617-21432 p)	54	Identities = 101/278 (36%), Positives = 164/278 (58%), Gaps = 6/278 (2%) gb AAA57943.1 (U18997) ORF_o290; Geneplot suggests frameshift linking to o267, not found [Escherichia coli] Length = 290
SeqID 274	SA-1132.1	Contig120 (21512-23416 p)	19	Identities = 76/206 (36%), Positives = 124/206 (59%), Gaps = 1/206 (0%) emb CAB01924.1 (Z79691) OrfA [Streptococcus pneumoniae] Length = 207
SeqID 275	SA-1134.1	Contig120 (23496-24497 p)	79	Identities = 222/333 (66%), Positives = 279/333 (83%) emb CAB01925.1 (Z79691) RegR [Streptococcus pneumoniae] Length = 333
SeqID 276	SA-1135.1	Contig120 (24568-25182 m)	87	Identities = 169/204 (82%), Positives = 192/204 (93%), Gaps = 1/204 (0%) emb CAC15392.1 (AJ278785) polypeptide deformylase [Streptococcus pneumoniae] Length = 203
SeqID 277	SA-1136.1	Contig120 (25248-25760 m)	42	Identities = 58/191 (30%), Positives = 98/191 (50%) sp P33023 YEIL_ECOLI_HYPOTHETICAL_25.3_KD_PROTEIN_IN_NFO-FRUA_INTERGENIC_REGION pir B64985 hypothetical 25.3 kD protein in nfo-fruA intergenic region - Escherichia coli (strain K-12) gb AAA60515.1 (U00007) yeiL [Escherichia coli] gb AAC75224.1 (AF000305) putative transcriptional regulator [Escherichia coli K12] prf 2014253BJ yeiL gene [Escherichia coli] Length = 219

SeqID 278	SA-1137.1	Contig120 (26052-27167 p)	43	Identities = 95/382 (24%), Positives = 174/382 (44%), Gaps = 30/382 (7%) pir D75094 transport protein, permease PAB0761 - Pyrococcus abyssi (strain Orsay) emb CAB50057.1 (AJ248286) TRANSPORT PROTEIN, permease [Pyrococcus abyssi] Length = 372
SeqID 279	SA-1138.1	Contig120 (27213-28196 m)	50	Identities = 101/327 (30%), Positives = 169/327 (50%), Gaps = 12/327 (3%) pir T50594 probable oxidoreductase [imported] - Streptomyces coelicolor emb CAB61731.1 (AL133220) putative oxidoreductase. [Streptomyces coelicolor A3(2)] Length = 351
SeqID 280	SA-1139.1	Contig120 (28344-30239 p)	88	Identities = 504/631 (79%), Positives = 563/631 (88%) emb CAC14579.1 (AJ249396) oligopeptidase [Streptococcus thermophilus] Length = 631
SeqID 281	SA-114.1	Contig131 (8878-9219 m)	57	Identities = 66/114 (57%), Positives = 85/114 (73%) spiP16963 YSA1_STRPY HYPOTHETICAL 16.6 KD PROTEIN IN SAGP 3 REGION pir JE0061 hypothetical 16.5K protein (SAGP 5 region) - Streptococcus pyogenes dbj BAA02939.1 (D13790) ORF2 [Streptococcus pyogenes] Length = 146
SeqID 282	SA-1140.1	Contig120 (30289-31242 m)	61	Identities = 133/309 (43%), Positives = 197/309 (63%), Gaps = 5/309 (1%) sp O35046 YOC_D_BACSU HYPOTHETICAL 36.4 KDA PROTEIN IN CSA-DES INTERGENIC REGION pir A69901 immunity to bacteriotoxins homolog yocD - Bacillus subtilis gb AAB84435.1 (AF027868) YocD [Bacillus subtilis] emb CAB13809.1 (Z99114) similar to immunity to bacteriotoxins [Bacillus subtilis] Length = 325
SeqID 283	SA-1141.1	Contig120 (31400-31645 p)	No Hits found	
SeqID 284	SA-1143.2	Contig120 (31746-33716 p)	39	Identities = 174/524 (33%), Positives = 275/524 (52%), Gaps = 13/524 (2%) pir E70040 conserved hypothetical protein yvGP - Bacillus subtilis emb CAB15347.1 (Z99121) similar to hypothetical proteins [Bacillus subtilis] Length = 670

SeqID 285	SA-1144.1	Contig120 (33993-35525 m)	86	Identities = 383/509 (75%), Positives = 441/509 (86%) dbj BAA25696.1 (AB010712) NADH oxidase/alkyl hydroperoxidase reductase [Streptococcus mutans] Length = 510
SeqID 286	SA-1145.1	Contig120 (35543-36103 m)	92	Identities = 168/186 (90%), Positives = 180/186 (96%) dbj BAA25695.1 (AB010712) alkyl hydroperoxidase [Streptococcus mutans] Length = 186
SeqID 287	SA-1146.1	Contig124 (43925-44143 m)	67	Identities = 36/72 (50%), Positives = 49/72 (68%) sp P54457 YQEL_BACSU HYPOTHETICAL 13.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION pir H69951 ybeB protein homolog yqeL - Bacillus subtilis dbj BAA12449.1 (D84432) YqeL [Bacillus subtilis] emb CAB14504.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 118
SeqID 288	SA-1148.1	Contig124 (43125-43859 m)	61	Identities = 86/242 (35%), Positives = 155/242 (63%), Gaps = 4/242 (1%) sp P54458 YQEM_BACSU HYPOTHETICAL 28.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION pir A69952 conserved hypothetical protein yqeM - Bacillus subtilis dbj BAA12450.1 (D84432) YqeM [Bacillus subtilis] emb CAB14503.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 247
SeqID 289	SA-1149.1	Contig124 (42851-43075 m)	No Hits found	
SeqID 290	SA-115.1	Contig131 (7849-8862 m)	86	Identities = 244/333 (73%), Positives = 294/333 (88%), Gaps = 1/333 (0%) pir T46742 ornithine carbamoyltransferase (EC 2.1.3.3) [validated] - Lactobacillus sakei emb CAA04683.1 (AJ001330) ornithine transcarbamoylase [Lactobacillus sakei] Length = 337
SeqID 291	SA-1150.1	Contig124 (41745-42854 m)	49	Identities = 131/367 (35%), Positives = 185/367 (49%), Gaps = 45/367 (12%) dbj BAB06304.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 416

SeqID 292	SA-1151.1	Contig124 (40939-41652 m)	56	Identities = 92/246 (37%), Positives = 143/246 (57%), Gaps = 14/246 (5%) pir C69661 transcription activator of multidrug-efflux transporter genes mta - Bacillus subtilis emb CAA69863.1 (Y08559) Unknown [Bacillus subtilis] emb CAB15677.1 (Z99122) transcriptional regulator [Bacillus subtilis] Length = 257
SeqID 293	SA-1152.1	Contig124 (40396-40800 m)	49	Identities = 56/129 (43%), Positives = 74/129 (56%), Gaps = 7/129 (5%) sp P54430 YRKC_BACSU HYPOTHETICAL 21.3 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION pir JA69976 hypothetical protein yrkC - Bacillus subtilis dbj BAA12358.1 (D84432) YrkC [Bacillus subtilis] emb CAB14597.1 (Z99117) yrkC [Bacillus subtilis] Length = 186
SeqID 294	SA-1153.1	Contig124 (39417-40130 p)	70	Identities = 136/230 (59%), Positives = 171/230 (74%) dbj BAB04811.1 (AP001510) glycerol uptake facilitator [Bacillus halodurans] Length = 276
SeqID 295	SA-1154.1	Contig124 (39033-39407 p)	63	Identities = 64/118 (54%), Positives = 85/118 (71%) dbj BAB07114.1 (AP001518) unknown conserved protein in others [Bacillus halodurans] Length = 128
SeqID 296	SA-1155.1	Contig124 (38455-39033 p)	68	Identities = 99/193 (51%), Positives = 138/193 (71%), Gaps = 2/193 (1%) dbj BAB07115.1 (AP001518) unknown conserved protein in others [Bacillus halodurans] Length = 196
SeqID 297	SA-1156.1	Contig124 (37428-38417 p)	76	Identities = 205/329 (62%), Positives = 261/329 (79%) dbj BAB07116.1 (AP001518) dihydroxyacetone kinase [Bacillus halodurans] Length = 330
SeqID 298	SA-1157.1	Contig124 (36718-37254 m)	44	Identities = 59/142 (41%), Positives = 81/142 (56%), Gaps = 5/142 (3%) dbj BAB07113.1 (AP001518) unknown [Bacillus halodurans] Length = 186
SeqID 299	SA-1158.1	Contig124 (35722-36708 m)	59	Identities = 142/285 (49%), Positives = 197/285 (68%), Gaps = 1/285 (0%) dbj BAB07112.1 (AP001518) dihydroxyacetone kinase [Bacillus halodurans] Length = 331

SeqID 300	SA-116.1	Contig131 (6359-7786 m)	72	Identities = 265/470 (56%), Positives = 347/470 (73%), Gaps = 3/470 (0%) emb CAA76779.1 (Y17554) permease [Bacillus licheniformis] Length = 468
SeqID 301	SA-1160.1	Contig124 (34888-35604 m)	72	Identities = 182/237 (76%), Positives = 201/237 (84%) sp P76351 YEEN_ECOLI_HYPOTHETICAL_25.9_KDA_PROTEIN_IN_AMN-CBL_INTERGENIC_REGION_pir A64963 conserved hypothetical protein b1983 - Escherichia coli gb AAC75047.1 (AE000290) orf, hypothetical protein [Escherichia coli K12] Length = 238
SeqID 302	SA-1161.1	Contig124 (34044-34733 m)	43	Identities = 72/178 (40%), Positives = 101/178 (56%), Gaps = 2/178 (1%) pir T36850 hypothetical protein SC135.37 - Streptomyces coelicolor emb CAA20826.1 (AL031541) hypothetical protein SC135.37 [Streptomyces coelicolor A3(2)] Length = 242
SeqID 303	SA-1162.1	Contig124 (33109-33942 m)	46	Identities = 83/264 (31%), Positives = 133/264 (49%), Gaps = 13/264 (4%) gb AAC45332.1 (U97348) basic surface protein [Lactobacillus fermentum] Length = 264
SeqID 304	SA-1163.1	Contig124 (32144-32968 m)	58	Identities = 110/284 (38%), Positives = 176/284 (61%), Gaps = 13/284 (4%) emb CAB59825.1 (AJ012388) hypothetical protein [Lactococcus lactis] Length = 287
SeqID 305	SA-1165.1	Contig124 (30634-32010 m)	46	Identities = 132/419 (31%), Positives = 212/419 (50%), Gaps = 14/419 (3%) pir E75327 ArgE/DapE/Acy1 family protein - Deinococcus radiodurans (strain R1) gb AAF11560.1 AE002038_9 (AE002038) ArgE/DapE/Acy1 family protein [Deinococcus radiodurans] Length = 463
SeqID 306	SA-1166.1	Contig124 (29833-30207 p)	38	Identities = 43/115 (37%), Positives = 52/115 (44%), Gaps = 2/115 (1%) pir H72617 hypothetical protein APE1401 - Aeropyrum pernix (strain K1) dbj BAA80398.1 (AP000061) 123aa long hypothetical protein [Aeropyrum pernix] Length = 123

SeqID 307	SA-1167.1	Contig124 (29571-30641 m)	71	Identities = 187/338 (55%), Positives = 256/338 (75%), Gaps = 12/338 (3%) emb[CAB59828.1] (AJ012388) hypothetical protein [Lactococcus lactis] Length = 368
SeqID 308	SA-1169.1	Contig124 (28882-29574 m)	73	Identities = 137/231 (59%), Positives = 171/231 (73%), Gaps = 1/231 (0%) emb[CAB59829.1] (AJ012388) hypothetical protein [Lactococcus lactis] Length = 231
SeqID 309	SA-117.1	Contig131 (5382-6338 m)	76	Identities = 199/311 (63%), Positives = 254/311 (80%), Gaps = 3/311 (0%) sp O53090 ARCC_LACSK CARBAMATE KINASE pir T46743 carbamate kinase (EC 2.7.2.2) [validated] - Lactobacillus sakei emb[CAA04684.1] (AJ001330) carbamate kinase [Lactobacillus sakei] Length = 314
SeqID 310	SA-1170.1	Contig124 (27778-28833 p)	72	Identities = 196/347 (56%), Positives = 257/347 (73%), Gaps = 1/347 (0%) pir C81088 alcohol dehydrogenase, zinc-containing NMB1395 [imported] - Neisseria meningitidis (group B strain MD58) gb AAF41759.1 (AE002488) alcohol dehydrogenase, zinc-containing [Neisseria meningitidis MC58] Length = 346
SeqID 311	SA-1171.2	Contig124 (26294-27664 p)	55	Identities = 182/453 (40%), Positives = 265/453 (58%), Gaps = 29/453 (6%) pir C69596 branched-chain amino acid transporter braB - Bacillus subtilis gb AAC00400.1 (AF008220) branch-chain amino acid transporter [Bacillus subtilis] emb[CAB14938.1] (Z99119) branched-chain amino acid transporter [Bacillus subtilis] Length = 445
SeqID 312	SA-1173.3	Contig136 (3489-3797 m)	No Hits found	
SeqID 313	SA-1174.1	Contig136 (3794-4186 m)	No Hits found	
SeqID 314	SA-1175.1	Contig136 (4104-4676 m)	No Hits found	
SeqID 315	SA-1176.1	Contig136 (4677-5165 m)	No Hits found	
SeqID 316	SA-1177.1	Contig136 (5180-5494 m)	No Hits found	
SeqID 317	SA-1178.1	Contig136 (5729-6937 m)	No Hits found	
SeqID 318	SA-1179.1	Contig136 (6986-7477 m)	No Hits found	

SeqID 319	SA-118.1	Contig131 (4247-5272 p)	78	Identities = 225/340 (66%), Positives = 271/340 (79%), Gaps = 3/340 (0%) sp Q46127 SYW_CLOLO TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS) gb AAC05711.1 (L49336) tryptophanyl-tRNA synthetase [Clostridium longisporum] Length = 341
SeqID 320	SA-1180.2	Contig136 (7491-11102 m)	52	Identities = 405/1293 (31%), Positives = 636/1293 (48%), Gaps = 175/1293 (13%) pir D71810 probable type II DNA modification enzyme (methyltransferase) - Helicobacter pylori (strain J99) gb AAD06987.1 (AE001563) putative TYPE II DNA MODIFICATION ENZYME (METHYLTRANSFERASE) [Helicobacter pylori J99] Length = 1252
SeqID 321	SA-1182.1	Contig136 (11439-12641 m)	40	Identities = 90/357 (25%), Positives = 162/357 (45%), Gaps = 33/357 (9%) ref NP_049990.1 orf359 gp [Streptococcus thermophilus bacteriophage Sfi21] emb CAA64931.1 (X95646) integrase [Streptococcus thermophilus bacteriophage Sfi21] gb AAC03454.1 (AF020798) integrase homolog [Streptococcus thermophilus bacteriophage TP-J34] gb AAD44095.1 AF115103_25 (AF115103) orf359 gp [Streptococcus thermophilus bacteriophage Sfi21] Length = 359
SeqID 322	SA-1183.1	Contig136 (12641-12952 m)	No Hits found	
SeqID 323	SA-1184.1	Contig136 (13608-14501 m)	74	Identities = 183/298 (61%), Positives = 234/298 (78%), Gaps = 1/298 (0%) sp P23496 LAXP_LACLA LACX PROTEIN, PLASMID pir D23696 lacX protein - Lactococcus lactis gb AAA25184.1 (M60447) ORF [Lactococcus lactis] Length = 299
SeqID 324	SA-1187.1	Contig136 (14601-16007 m)	96	Identities = 442/468 (94%), Positives = 459/468 (97%) gb AAA26949.1 (M19454) phospho-beta-D-galactosidase (EC 3.2.1.85) [Lactococcus lactis subsp. cremoris] Length = 477
SeqID 325	SA-1188.1	Contig136 (16587-17003 p)	No Hits found	

SeqID 326	SA-1189.1	Contig136 (16094-17800 m)	94	<p>Identities = 520/568 (91%), Positives = 547/568 (95%), Gaps = 1/568 (0%) sp P23531 PTLB_LACLA PTS SYSTEM, LACTOSE-SPECIFIC IIBC COMPONENT (EIIBC-LAC) (LACTOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EI-I-LAC) pir B23696 phosphotransferase system enzyme II (EC 2.7.1.69) - Lactococcus lactis gb AAA25182.1 (M60447) enzyme III [Lactococcus lactis] Length = 568</p>
SeqID 327	SA-119.1	Contig131 (4008-4139 p)	No Hits found	
SeqID 328	SA-1190.1	Contig136 (17800-18117 m)	78	<p>Identities = 89/105 (84%), Positives = 97/105 (91%) sp P23532 PTLA_LACLA PTS SYSTEM, LACTOSE-SPECIFIC IIA COMPONENT (EIIA-LAC) (LACTOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EI-I-LAC) pir A23696 phosphotransferase system enzyme III (EC 2.7.-) - Lactococcus lactis pdb 1E2A A Chain A, Enzyme IIA From The Lactose Specific Pts From Lactococcus Lactis pdb 1E2A B Chain B, Enzyme IIA From The Lactose Specific Pts From Lactococcus Lactis pdb 1E2A C Chain C, Enzyme IIA From The Lactose Specific Pts From Lactococcus Lactis gb AAA25181.1 (M60447) enzyme III [Lactococcus lactis] Length = 105</p>
SeqID 329	SA-1191.1	Contig136 (18146-18979 m)	56	<p>Identities = 104/278 (37%), Positives = 169/278 (60%), Gaps = 1/278 (0%) sp P24401 LACT_LACCA TRANSCRIPTION ANTIMINATOR LACT emb CAB02555.1 (Z80834) Lact [Lactobacillus casei] gb AAB49331.1 (U21391) Lact [Lactobacillus casei] Length = 292</p>

SeqID 330	SA-1192.1	Contig136 (19372-20352 m)	92	<p>Identities = 279/326 (85%), Positives = 308/326 (93%) sp P26593 LACD_LACLA_TAGATOSE_1,6-DIPHOSPHATE ALDOLASE (TAGATOSE-BISPHOSPHATE ALDOLASE) (D-TAGATOSE-1,6-BISPHOSPHATE ALDOLASE) pir D39778 LacD tagatose-1,6-diphosphate aldolase - Lactococcus lactis gb AAA25171.1 (M65190) lacD [Lactococcus lactis] gb AAA25180.1 (M60447) tagatose 1,6-diP aldolase [Lactococcus lactis] Length = 326</p>
SeqID 331	SA-1193.1	Contig136 (20357-21286 m)	78	<p>Identities = 207/310 (66%), Positives = 245/310 (78%), Gaps = 1/310 (0%) sp P23391 LACC_LACLA_TAGATOSE-6- PHOSPHATE KINASE (PHOSPHOTAGATOKINASE) pir C39778 tagatose-6-phosphate kinase (EC 2.7.1.-) LacC - Lactococcus lactis gb AAA25170.1 (M65190) lacC [Lactococcus lactis] gb AAA25179.1 (M60447) tagatose 6-P kinase [Lactococcus lactis] Length = 310</p>
SeqID 332	SA-1194.3	Contig136 (21299-21814 m)	85	<p>Identities = 142/171 (83%), Positives = 156/171 (91%) sp P23495 LACB_LACLA_GALACTOSE-6-PHOSPHATE ISOMERASE LACB SUBUNIT pir B39778 galactose-6- phosphate isomerase (EC 5.3.1.-) component LacB - Lactococcus lactis gb AAA25169.1 (M65190) lacB [Lactococcus lactis] gb AAA25178.1 (M60447) galactose 6-P isomerase [Lactococcus lactis] Length = 171</p>
SeqID 333	SA-1195.3	Contig136 (21831-22256 m)	76	<p>Identities = 120/141 (85%), Positives = 134/141 (94%) sp P23494 LACA_LACLA_GALACTOSE-6-PHOSPHATE ISOMERASE LACA SUBUNIT pir A39778 galactose-6- phosphate isomerase (EC 5.3.1.-) component LacA - Lactococcus lactis gb AAA25168.1 (M65190) lacA [Lactococcus lactis] gb AAA25177.1 (M60447) galactose 6-P isomerase [Lactococcus lactis] Length = 141</p>
SeqID 334	SA-1196.2	Contig111 (10300-10728 p)	No Hits found	

SeqID 335	SA-1197.1	Contig111 (9929-10219 p)	49	Identities = 35/91 (38%), Positives = 51/91 (55%) splP34159 YHS1_CLOAB HYPOTHETICAL 11.0 KD PROTEIN IN HSP18 3 REGION (ORFA1) emb CAA46375.1 (X65276) ORFA1 [Clostridium acetobutylicum] Length = 96
SeqID 336	SA-1198.1	Contig111 (9134-9421 m)	No Hits found	
SeqID 337	SA-1199.1	Contig111 (8809-9144 m)	No Hits found	
SeqID 338	SA-120.1	Contig110 (1487-1594 p)	No Hits found	
SeqID 339	SA-1200.1	Contig111 (8448-8687 m)	No Hits found	
SeqID 340	SA-1201.1	Contig111 (8034-8504 p)	No Hits found	
SeqID 341	SA-1202.1	Contig111 (6984-7439 m)	34	Identities = 36/89 (40%), Positives = 52/89 (57%), Gaps = 3/89 (3%) pir T39903 serine-rich protein - fission yeast (Schizosaccharomyces pombe) emb CAA22127.1 (AL033534) hypothetical serine-rich secreted protein [Schizosaccharomyces pombe] Length = 534
SeqID 342	SA-1203.1	Contig111 (6610-7866 p)	59	Identities = 171/402 (42%), Positives = 248/402 (61%), Gaps = 46/402 (11%) pir A33952 58K mobilization protein - Streptococcus pneumoniae plasmids gb AA25387.1 (M28538) mobilization peptide [Plasmid pMV158] Length = 494
SeqID 343	SA-1205.1	Contig111 (5859-6293 p)	No Hits found	
SeqID 344	SA-1207.1	Contig111 (4967-5755 p)	51	Identities = 90/264 (34%), Positives = 139/264 (52%), Gaps = 22/264 (8%) prf 1405330A repM gene [Staphylococcus aureus] Length = 314
SeqID 345	SA-1208.1	Contig111 (4026-4667 p)	No Hits found	
SeqID 346	SA-1209.1	Contig111 (3684-4022 p)	49	Identities = 34/102 (33%), Positives = 56/102 (54%) gb AAD41248.1 AF106927_1 (AF106927) unknown [Streptococcus suis] Length = 272
SeqID 347	SA-121.1	Contig131 (3142-3447 m)	No Hits found	
SeqID 348	SA-1210.1	Contig111 (3409-3690 p)	No Hits found	
SeqID 349	SA-1211.1	Contig111 (2975-3268 p)	No Hits found	
SeqID 350	SA-1212.1	Contig111 (2259-2810 m)	15	Identities = 23/63 (36%), Positives = 36/63 (56%) dbj BAB05162.1 (AP001512) transcriptional regulator [Bacillus halodurans] Length = 107

SeqID 351	SA-1213.1	Contig111 (1036-2190 m)	53	Identities = 128/386 (33%), Positives = 208/386 (53%), Gaps = 18/386 (4%) gb AAG29618.1 (AF217235) integrase-like protein [Staphylococcus aureus] Length = 378
SeqID 352	SA-1214.1	Contig111 (516-908 p)	73	Identities = 90/129 (69%), Positives = 106/129 (81%) sp P07842 RS9_BACST_30S RIBOSOMAL PROTEIN S9 (BS10) pir R3BS9 ribosomal protein S9 - Bacillus stearothermophilus Length = 129
SeqID 353	SA-1215.1	Contig111 (49-495 p)	68	Identities = 89/144 (61%), Positives = 113/144 (77%) dbj BAB03887.1 (AP001507) ribosomal protein L13 [Bacillus halodurans] Length = 145
SeqID 354	SA-1216.2	Contig130 (62901-63572 p)	24	Identities = 27/119 (22%), Positives = 54/119 (44%), Gaps = 10/119 (8%) emb CAC12089.1 (AL445066) NADH dehydrogenase, chain M related protein [Thermoplasma acidophilum] Length = 503
SeqID 355	SA-1217.1	Contig130 (62156-62926 p)	66	Identities = 121/249 (48%), Positives = 172/249 (68%) pir H70023 N-acetyl-glucosamine catabolism homolog yutF - Bacillus subtilis emb CAB15219.1 (Z99120) similar to N-acetyl-glucosamine catabolism [Bacillus subtilis] Length = 256
SeqID 356	SA-1218.1	Contig130 (61418-62155 p)	42	Identities = 59/248 (23%), Positives = 104/248 (41%), Gaps = 30/248 (12%) pir S51698 oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) - Arabidopsis thaliana pir S69197 oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) (clone TE 1-7) - Arabidopsis thaliana emb CAA85389.1 (Z36912) acyl-(acyl carrier protein) thioesterase [Arabidopsis thaliana] dbj BAB02069.1 (AB026647) acyl carrier protein thioesterase [Arabidopsis thaliana] Length = 362
SeqID 357	SA-1219.1	Contig130 (60284-61414 p)	65	Identities = 173/375 (46%), Positives = 248/375 (66%), Gaps = 5/375 (1%) dbj BAB05062.1 (AP001511) coproporphyrinogen III oxidase [Bacillus halodurans] Length = 385
SeqID 358	SA-1220.1	Contig130 (59812-60192 p)	No Hits found	

SeqID 359	SA-1221.1	Contig130 (59416-59787 p)	26	<p>Identities = 20/68 (29%), Positives = 35/68 (51%) sp P24252 YBGA_ECOLI HYPOTHETICAL 20.2 KDA PROTEIN IN RHSC-PHRB INTERGENIC REGION (TKP) (ORF169) pir B64806 ybgA protein - Escherichia coli gb AAA24387.1 (K01299) unidentified protein ORF169 [Escherichia coli] gb AAC73801.1 (AE000174) orf, hypothetical protein [Escherichia coli K12] dbj BAA35366.1 (D90709) Hypothetical 20.2 kd protein in phrB 5 region (tkp) (orf169). [Escherichia coli] dbj BAA35371.1 (D90710) Hypothetical 20.2 kd protein in phrB 5 region (tkp) (orf169). [Escherichia coli] Length = 169</p>
SeqID 360	SA-1222.1	Contig130 (57941-59293 p)	77	<p>Identities = 284/451 (62%), Positives = 353/451 (77%), Gaps = 4/451 (0%) pir B69745 phosphoglucosyltransferase (glycolysis) homolog ybbT - Bacillus subtilis emb CAB11953.1 (Z99104) similar to phosphoglucosyltransferase (glycolysis) [Bacillus subtilis] emb CAB11970.1 (Z99105) similar to phosphoglucosyltransferase (glycolysis) [Bacillus subtilis] Length = 448</p>
SeqID 361	SA-1223.1	Contig130 (57018-57887 p)	50	<p>Identities = 90/324 (27%), Positives = 167/324 (50%), Gaps = 18/324 (5%) pir A69745 hypothetical protein ybbR - Bacillus subtilis emb CAB11952.1 (Z99104) ybbR [Bacillus subtilis] emb CAB11969.1 (Z99105) ybbR [Bacillus subtilis] Length = 483</p>
SeqID 362	SA-1225.1	Contig130 (56011-56925 p)	60	<p>Identities = 125/253 (49%), Positives = 186/253 (73%), Gaps = 5/253 (1%) pir H69744 conserved hypothetical protein ybbP - Bacillus subtilis dbj BAA19509.1 (AB002150) YbbP [Bacillus subtilis] emb CAB11951.1 (Z99104) alternate gene name: ybbQ~similar to hypothetical proteins [Bacillus subtilis] emb CAB11968.1 (Z99105) alternate gene name: ybbQ~similar to hypothetical proteins [Bacillus subtilis] Length = 273</p>

SeqID 363	SA-1226.1	Contig130 (54591-55910 m)	50	Identities = 144/442 (32%), Positives = 229/442 (51%), Gaps = 17/442 (3%) pir T31440 UDP-N-acetyluramyl tripeptide synthetase homolog murC - Heliobacillus mobilis gb AAC84012.1 (AF080002) UDP-N-acetyluramyl tripeptide synthetase MurC [Heliobacillus mobilis] Length = 455
SeqID 364	SA-1227.1	Contig130 (53806-54477 m)	47	Identities = 89/250 (35%), Positives = 129/250 (51%), Gaps = 9/250 (3%) pir T31439 probable cobyric acid synthase CobQ - Heliobacillus mobilis gb AAC84011.1 (AF080002) cobyric acid synthase CobQ [Heliobacillus mobilis] Length = 252
SeqID 365	SA-1228.1	Contig130 (52710-53699 p)	63	Identities = 153/316 (48%), Positives = 212/316 (66%), Gaps = 3/316 (0%) dbj BAB04402.1 (AP001509) lipoate-protein ligase [Bacillus halodurans] Length = 330
SeqID 366	SA-123.1	Contig131 (1559-1945 p)	79	Identities = 95/128 (74%), Positives = 105/128 (81%), Gaps = 8/128 (6%) sp P20277 RL17_BACSU 50S RIBOSOMAL PROTEIN L17 (BL15) (BL21) pir F32307 ribosomal protein L17 - Bacillus subtilis gb AAA22218.1 (M26414) ribosomal protein L17 [Bacillus subtilis] gb AAB06827.1 (L47971) ribosomal protein L17 [Bacillus subtilis] emb CAB11920.1 (Z99104) ribosomal protein L17 (BL15) [Bacillus subtilis] Length = 120
SeqID 367	SA-1230.1	Contig130 (50855-52612 p)	57	Identities = 229/589 (38%), Positives = 339/589 (56%), Gaps = 25/589 (4%) pir J40794 dihydrolipoamide dehydrogenase (EC 1.8.1.4) [validated] - Clostridium magnum gb AAA21748.1 (L31844) dihydrolipoamide dehydrogenase [Clostridium magnum] Length = 578
SeqID 368	SA-1231.1	Contig130 (49407-50795 p)	57	Identities = 187/462 (40%), Positives = 267/462 (57%), Gaps = 26/462 (5%) dbj BAB04497.1 (AP001509) dihydrolipoamide S-acetyltransferase [Bacillus halodurans] Length = 436

SeqID 369	SA-1232.1	Contig130 (48282-49280 p)	72	Identities = 189/319 (59%), Positives = 249/319 (77%), Gaps = 1/319 (0%) dbj BAB04496.1 (AP001509) acetoin dehydrogenase (TPP-dependent) beta chain [Bacillus halodurans] Length = 327
SeqID 370	SA-1233.1	Contig130 (48230-48595 m)	40	Identities = 45/97 (46%), Positives = 50/97 (51%) pir G72548 hypothetical protein APE1675 - Aeropyrum pernix (strain K1) dbj BAA80676.1 (AP000062) 155aa long hypothetical protein [Aeropyrum pernix] Length = 155
SeqID 371	SA-1234.1	Contig130 (47239-48207 p)	64	Identities = 148/317 (46%), Positives = 214/317 (66%), Gaps = 1/317 (0%) dbj BAB04495.1 (AP001509) acetoin dehydrogenase (TPP-dependent) alpha chain [Bacillus halodurans] Length = 326
SeqID 372	SA-1236.3	Contig130 (45180-47090 p)	75	Identities = 397/641 (61%), Positives = 488/641 (75%), Gaps = 17/641 (2%) gb AAK05160.1 AE006339_3 (AE006339) ABC transporter ATP binding protein [Lactococcus lactis subsp. lactis] Length = 635
SeqID 373	SA-1238.1	Contig135 (95462-95596 p)	No Hits found	
SeqID 374	SA-1239.1	Contig135 (94858-95316 p)	No Hits found	
SeqID 375	SA-1240.1	Contig135 (94570-94797 p)	No Hits found	
SeqID 376	SA-1241.1	Contig135 (94477-94533 p)	No Hits found	
SeqID 377	SA-1242.1	Contig135 (93779-94036 m)	No Hits found	
SeqID 378	SA-1243.1	Contig135 (93763-94326 p)	No Hits found	
SeqID 379	SA-1244.1	Contig135 (93063-93815 p)	No Hits found	
SeqID 380	SA-1245.1	Contig135 (92529-93014 p)	No Hits found	
SeqID 381	SA-1246.1	Contig135 (92196-92423 p)	No Hits found	
SeqID 382	SA-1247.1	Contig135 (89123-91921 p)	22	Identities = 157/406 (38%), Positives = 216/406 (52%), Gaps = 28/406 (6%) gb AAB30711.1 (S70345) SpaA=endocarditis immunodominant antigen [Streptococcus sobrinus, MUCOB 263, Peptide, 1566 aa] Length = 1566

SeqID 383	SA-1248.1	Contig135 (88346-89056 p)	53	Identities = 85/268 (31%), Positives = 129/268 (47%), Gaps = 35/268 (13%) pir S45091 hypothetical protein iota - Streptococcus pyogenes plasmid pDB101 pir S68125 hypothetical protein iota - Streptococcus pyogenes plasmid pSM19035 emb CAA47097.1 (X66468) orf iota [Streptococcus pyogenes] Length = 268
SeqID 384	SA-1249.1	Contig135 (88324-88794 m)	30	Identities = 37/104 (35%), Positives = 48/104 (45%), Gaps = 5/104 (4%) ref XP_007094.1 collagen, type IV, alpha 1 [Homo sapiens] Length = 1669
SeqID 385	SA-125.1	Contig131 (606-1544 p)	79	Identities = 190/314 (60%), Positives = 249/314 (78%), Gaps = 2/314 (0%) sp O50634 RPOA_BACHD DNA-DIRECTED RNA POLYMERASE ALPHA CHAIN (TRANSCRIPTASE ALPHA CHAIN) (RNA POLYMERASE ALPHA SUBUNIT) pir T44410 DNA-directed RNA polymerase (EC 2.7.7.6) alpha chain rpoA [imported] - Bacillus halodurans dbj BAA24194.1 (AB010082) RNA polymerase alpha subunit [Bacillus halodurans] dbj BAA75298.1 (AB017508) rpoA homologue (identity of 85 to B. subtilis) [Bacillus halodurans] dbj BAB03881.1 (AP001507) DNA-directed RNA polymerase alpha subunit [Bacillus halodurans] Length = 314
SeqID 386	SA-1250.1	Contig135 (86100-88331 p)	45	Identities = 185/839 (22%), Positives = 342/839 (40%), Gaps = 124/839 (14%) pir G41662 130K surface exclusion protein Sec10 precursor - Enterococcus faecalis plasmid pCF10 gb AAA65847.1 (M64978) surface exclusion protein [Plasmid pCF10] Length = 891
SeqID 387	SA-1251.1	Contig135 (85784-86083 p)	No Hits found	
SeqID 388	SA-1252.1	Contig135 (85204-85401 p)	60	Identities = 29/58 (50%), Positives = 41/58 (70%) ref NP_049430.1 unknown [Streptococcus thermophilus bacteriophage DT1] gb AAD21918.1 (AF085222) unknown [Streptococcus thermophilus bacteriophage DT1] Length = 165
SeqID 389	SA-1253.1	Contig135 (85022-85207 p)	No Hits found	

SeqID 390	SA-1254.1	Contig135 (83986-85020 p)	36	Identities = 95/192 (49%), Positives = 128/192 (66%), Gaps = 7/192 (3%) ref NP_038742.1 IFN-response element binding factor 1 [Mus musculus] sp P22560 RBF1_MOUSE IFN-RESPONSE BINDING FACTOR 1 (IREBF-1) pir A38558 interferon response element-binding factor 1 - mouse gb AAA37884.1 (M55290) IFN-response element binding factor 1 [Mus musculus] Length = 311
SeqID 391	SA-1255.1	Contig135 (83829-83993 p)	No Hits found	
SeqID 392	SA-1256.1	Contig135 (83345-83722 p)	No Hits found	
SeqID 393	SA-1257.1	Contig135 (82756-83013 p)	50	Identities = 28/81 (34%), Positives = 46/81 (56%), Gaps = 3/81 (3%) sp P21318 YR7D_ECOLI_HYPOTHETICAL 11.0 KD PROTEIN (ORFD) (RETRON EC67) pir JQ0859 hypothetical 11K protein - Escherichia coli retron Ec67 gb AAA23395.1 (M55249) unknown [Escherichia coli] Length = 100
SeqID 394	SA-1258.1	Contig135 (82201-82572 p)	No Hits found	
SeqID 395	SA-1261.1	Contig135 (81104-82087 p)	82	Identities = 244/321 (76%), Positives = 286/321 (89%), Gaps = 2/321 (0%) pir C70015 probable GMP reductase (EC 1.6.6.8) yumD - Bacillus subtilis emb CAB07955.1 (Z93939) unknown [Bacillus subtilis] emb CAB15203.1 (Z99120) similar to GMP reductase [Bacillus subtilis] Length = 326
SeqID 396	SA-1265.2	Contig125 (17134-18261 m)	75	Identities = 209/376 (55%), Positives = 286/376 (75%), Gaps = 3/376 (0%) gb AAK06013.1 AE006422_2 (AE006422) Na+/H+ antiporter [Lactococcus lactis subsp. lactis] Length = 379
SeqID 397	SA-1266.1	Contig125 (16108-17070 p)	45	Identities = 86/319 (26%), Positives = 147/319 (45%), Gaps = 34/319 (10%) sp P31847 YPUA_BACSU_HYPOTHETICAL 31.3 KD PROTEIN IN LYSA-PPIB INTERGENIC REGION (ORFX19) pir JU0473 ypuA protein - Bacillus subtilis gb AAA67474.1 (L09228) ORFX19 [Bacillus subtilis] emb CAB14269.1 (Z99116) ypuA [Bacillus subtilis] Length = 290

SeqID 398	SA-1267.1	Contig125 (15541-15954 p)	64	Identities = 76/139 (54%), Positives = 92/139 (65%), Gaps = 12/139 (8%) sp P45871 YWKD_BACSU HYPOTHETICAL 14.8 KD PROTEIN IN TDK-PRFA INTERGENIC REGION pir S55436 conserved hypothetical protein ywkD - Bacillus subtilis emb CAA89883.1 (Z49782) unknown [Bacillus subtilis] emb CAB15719.1 (Z99122) similar to hypothetical proteins [Bacillus subtilis] Length = 128
SeqID 399	SA-1268.1	Contig125 (14782-15525 p)	35	Identities = 65/126 (51%), Positives = 89/126 (70%), Gaps = 1/126 (0%) gb AAA73122.1 (M77279) alpha-amylase [unidentified cloning vector] Length = 137
SeqID 400	SA-127.1	Contig131 (173-556 p)	87	Identities = 107/123 (86%), Positives = 115/123 (92%) sp P04969 RS11_BACSU 30S RIBOSOMAL PROTEIN S11 (BS11) pir R3BSS1 ribosomal protein S11 - Bacillus subtilis gb AAA22216.1 (M26414) ribosomal protein S11 [Bacillus subtilis] gb AAA22707.1 (M13957) ribosomal protein S11 [Bacillus subtilis] gb AAB06825.1 (L47971) ribosomal protein S11 [Bacillus subtilis] emb CAB11918.1 (Z99104) ribosomal protein S11 (BS11) [Bacillus subtilis] Length = 131
SeqID 401	SA-1271.1	Contig125 (12316-14775 p)	92	Identities = 712/819 (86%), Positives = 769/819 (92%) gb AAF63266.1 (AF220945) DNA gyrase A subunit [Streptococcus pyogenes] Length = 828
SeqID 402	SA-1273.1	Contig125 (11088-12077 m)	90	Identities = 290/329 (88%), Positives = 313/329 (94%), Gaps = 1/329 (0%) sp O33734 LDH_STRPN L-LACTATE DEHYDROGENASE emb CAA04010.1 (AJ000336) L-lactate dehydrogenase [Streptococcus pneumoniae] Length = 328
SeqID 403	SA-1274.1	Contig125 (9598-10968 p)	88	Identities = 363/458 (79%), Positives = 408/458 (88%), Gaps = 3/458 (0%) gb AAC26485.2 (AF014458) NADH oxidase [Streptococcus pneumoniae] Length = 459

SeqID 404	SA-1275.1	Contig125 (8436-9392 p)	72	<p>Identities = 172/318 (54%), Positives = 234/318 (73%) pir F70009 conserved hypothetical protein yufQ - Bacillus subtilis emb CAB07939.1 (Z93937) unknown [Bacillus subtilis] emb CAB15146.1 (Z99120) similar to hypothetical proteins [Bacillus subtilis] Length = 319</p>
SeqID 405	SA-1278.1	Contig125 (7373-8434 p)	60	<p>Identities = 154/349 (44%), Positives = 220/349 (62%), Gaps = 6/349 (1%) pir E70009 conserved hypothetical protein yufP - Bacillus subtilis emb CAB07938.1 (Z93937) unknown [Bacillus subtilis] emb CAB15145.1 (Z99120) similar to hypothetical proteins [Bacillus subtilis] Length = 348</p>
SeqID 406	SA-1279.1	Contig125 (5845-7380 p)	76	<p>Identities = 311/497 (62%), Positives = 396/497 (79%), Gaps = 1/497 (0%) pir D70009 probable ABC transporter yufO - Bacillus subtilis emb CAB07937.1 (Z93937) unknown [Bacillus subtilis] emb CAB15144.1 (Z99120) similar to ABC transporter (ATP- binding protein) [Bacillus subtilis] Length = 510</p>
SeqID 407	SA-128.1	Contig131 (3-155 p)	86	<p>Identities = 41/50 (82%), Positives = 44/50 (88%) sp P15757 RS13_BACST 30S RIBOSOMAL PROTEIN S13 pir R3BS3F ribosomal protein S13 - Bacillus stearothermophilus Length = 119</p>
SeqID 408	SA-1280.1	Contig125 (4651-5700 p)	61	<p>Identities = 164/337 (48%), Positives = 225/337 (66%), Gaps = 10/337 (2%) sp O05252 YUFN_BACSU HYPOTHETICAL LIPOPROTEIN YUFN PRECURSOR pir C70009 ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis emb CAB07936.1 (Z93937) unknown [Bacillus subtilis] emb CAB15143.1 (Z99120) similar to ABC transporter (lipoprotein) [Bacillus subtilis] Length = 350</p>
SeqID 409	SA-1281.1	Contig125 (4197-4586 p)	52	<p>Identities = 66/114 (57%), Positives = 81/114 (70%) emb CAB51906.1 (AJ237978) cytidine deaminase [Bacillus psychrophilus] Length = 136</p>

SeqID 410	SA-1282.1	Contig125 (3281-3871 p)	69	Identities = 90/201 (44%), Positives = 144/201 (70%), Gaps = 5/201 (2%) sp P37872 YBX_BACSU HYPOTHETICAL 22.5 KDA PROTEIN IN RPLL-RPOB INTERGENIC REGION (P23) (ORF23) pir F69751 conserved hypothetical protein yxbB - Bacillus subtilis gb AAB00971.1 (L24376) hypothetical protein [Bacillus subtilis] emb CAB11882.1 (Z99104) alternate gene name: ybaA~similar to hypothetical proteins [Bacillus subtilis] Length = 201
SeqID 411	SA-1283.1	Contig125 (2255-3175 m)	65	Identities = 140/307 (45%), Positives = 203/307 (65%), Gaps = 5/307 (1%) sp Q9K8X7 COAA_BACHD PANTOTHENATE KINASE (PANTOTHENIC ACID KINASE) dbj BAB06594.1 (AP001516) pantothenate kinase [Bacillus halodurans] Length = 316
SeqID 412	SA-1284.1	Contig125 (1953-2195 m)	69	Identities = 47/86 (54%), Positives = 59/86 (67%), Gaps = (8%) dbj BAB05058.1 (AP001511) ribosomal protein S20 (BS20) [Bacillus halodurans] Length = 91
SeqID 413	SA-1285.1	Contig125 (1035-1868 p)	66	Identities = 170/269 (63%), Positives = 203/269 (75%), Gaps = 2/269 (0%) gb AAC35851.1 (AF086736) amino acid-binding protein Abp [Streptococcus uberis] Length = 277
SeqID 414	SA-1286.1	Contig125 (392-1021 p)	63	Identities = 94/210 (44%), Positives = 145/210 (68%), Gaps = 3/210 (1%) gb AAB49429.1 (U73111) glutamine transport ATP-binding protein GLNQ [Salmonella typhimurium] Length = 240
SeqID 415	SA-1287.1	Contig125 (196-381 p)	64	Identities = 54/125 (43%), Positives = 82/125 (65%), Gaps = 1/125 (0%) dbj BAA98402.1 (AP002545) ABC amino acid transporter permease [Chlamydia pneumoniae] Length = 217
SeqID 416	SA-1288.3	Contig139 (155214-156659 m)	21	Identities = 60/268 (22%), Positives = 107/268 (39%), Gaps = 42/268 (15%) gb AAG44891.1 (AF286595_1) (AF286595) flavin-containing monooxygenase FMO3 [Rattus norvegicus] Length = 531

SeqID 417	SA-1289.1	Contig139 (154486-155157 p)	66	Identities = 116/213 (54%), Positives = 152/213 (70%) gb AAD47593.1 AF140784_2 (AF140784) Vexp2 [Streptococcus pneumoniae] Length = 215
SeqID 418	SA-1290.1	Contig139 (154080-154274 p)	No Hits found	
SeqID 419	SA-1291.1	Contig139 (153049-154071 p)	45	Identities = 100/348 (28%), Positives = 158/348 (44%), Gaps = 17/348 (4%) gb AAD47594.1 AF140784_3 (AF140784) Vexp3 [Streptococcus pneumoniae] Length = 459
SeqID 420	SA-1292.1	Contig139 (152386-153039 p)	42	Identities = 61/210 (29%), Positives = 105/210 (49%), Gaps = 18/210 (8%) pir F69762 transporter homolog ycll - Bacillus subtilis dbj BAA09006.1 (D50453) homologue of hypothetical protein in a rapamycin synthesis gene cluster of Streptomyces hygroscopicus [Bacillus subtilis] emb CAB12182.1 (Z99106) similar to transporter [Bacillus subtilis] Length = 486
SeqID 421	SA-1293.1	Contig139 (151707-152243 p)	83	Identities = 175/178 (98%), Positives = 176/178 (98%) gb AAF01071.1 (AF157015) CylK [Streptococcus agalactiae] Length = 178
SeqID 422	SA-1295.1	Contig139 (150488-151699 p)	94	Identities = 396/403 (98%), Positives = 400/403 (98%) gb AAF01070.1 (AF157015) CylJ [Streptococcus agalactiae] Length = 403
SeqID 423	SA-1296.1	Contig139 (148288-150483 p)	95	Identities = 730/731 (99%), Positives = 731/731 (99%) gb AAF89495.1 (AF093787) CylI [Streptococcus agalactiae] Length = 731
SeqID 424	SA-1297.1	Contig139 (147338-148291 p)	99	Identities = 317/317 (100%), Positives = 317/317 (100%) gb AAF89494.1 (AF093787) CylF [Streptococcus agalactiae] Length = 317
SeqID 425	SA-1298.1	Contig139 (145338-147341 p)	99	Identities = 666/667 (99%), Positives = 667/667 (99%) gb AAD32040.1 AF093787_8 (AF093787) CylE [Streptococcus agalactiae] Length = 667
SeqID 426	SA-1299.1	Contig139 (144463-145341 p)	94	Identities = 291/292 (99%), Positives = 292/292 (99%) gb AAD32039.1 AF093787_7 (AF093787) ABC transporter homolog CylB [Streptococcus agalactiae] Length = 292
SeqID 427	SA-13.1	Contig137 (32737-32958 p)	No Hits found	

SeqID 428	SA-130.1	Contig138 (96433-96687 p)	No Hits found	
SeqID 429	SA-1300.1	Contig139 (143541-144470 p)	92	Identities = 308/309 (99%), Positives = 308/309 (99%) gb AAD32038.1 AF093787_6 (AF093787) ABC transporter homolog CylA [Streptococcus agalactiae] Length = 309
SeqID 430	SA-1301.1	Contig139 (143075-143551 p)	75	Identities = 120/120 (100%), Positives = 120/120 (100%) gb AAD32037.1 AF093787_5 (AF093787) CylZ [Streptococcus agalactiae] Length = 131
SeqID 431	SA-1302.1	Contig139 (142786-143091 p)	98	Identities = 101/101 (100%), Positives = 101/101 (100%) gb AAD32036.1 AF093787_4 (AF093787) acyl carrier protein homolog AcpC [Streptococcus agalactiae] Length = 101
SeqID 432	SA-1303.1	Contig139 (142071-142793 p)	94	Identities = 239/240 (99%), Positives = 240/240 (99%) gb AAD32035.1 AF093787_3 (AF093787) CylG [Streptococcus agalactiae] Length = 240
SeqID 433	SA-1304.2	Contig139 (141226-142074 p)	99	Identities = 282/282 (100%), Positives = 282/282 (100%) gb AAD32034.1 AF093787_2 (AF093787) CylD [Streptococcus agalactiae] Length = 282
SeqID 434	SA-1305.2	Contig139 (140921-141226 p)	93	Identities = 101/101 (100%), Positives = 101/101 (100%) gb AAD32033.1 AF093787_1 (AF093787) CylX [Streptococcus agalactiae] Length = 101
SeqID 435	SA-1308.2	Contig128 (27763-28422 p)	40	Identities = 54/181 (29%), Positives = 95/181 (51%), Gaps = 1/181 (0%) emb CAA65740.1 (X97014) PrfA [Listeria seeligeri] Length = 237
SeqID 436	SA-1309.1	Contig128 (28446-30731 p)	65	Identities = 386/767 (50%), Positives = 502/767 (65%), Gaps = 21/767 (2%) sp P22093 PEPX_LACLC XAA-PRO DIPEPTIDYL-PEPTIDASE (X-PRO DIPEPTIDYL-PEPTIDASE) (X-PROLYL-DIPEPTIDYL AMINOPEPTIDASE) (X-PDAP) pir A43747 X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Lactococcus lactis subsp. cremoris gb AAA25232.1 (M58315) dipeptidyl peptidase IV [Lactococcus lactis] Length = 763
SeqID 437	SA-1310.1	Contig128 (30735-31094 p)	No Hits found	

SeqID 438	SA-1311.1	Contig128 (31140-32120 p)	56	<p>Identities = 101/318 (31%), Positives = 184/318 (57%), Gaps = 5/318 (1%) sp P31114 HEP2_BACSU HEPTAPRENYL DIPHOSPHATE SYNTHASE COMPONENT II (HEPPP SYNTHASE SUBUNIT 2) (SPORE GERMINATION PROTEIN C3) pir E69630 heptaprenyl diphosphate synthase component II gerCC - Bacillus subtilis gb AAA20856.1 (M80245) GerC3 [Bacillus subtilis] emb CAB14190.1 (Z99115) heptaprenyl diphosphate synthase component II [Bacillus subtilis] Length = 348</p>
SeqID 439	SA-1312.1	Contig128 (32205-33953 m)	66	<p>Identities = 264/577 (45%), Positives = 391/577 (67%), Gaps = 14/577 (2%) sp P94367 CYDD_BACSU TRANSPORT ATP-BINDING PROTEIN CYDD pir D69611 ABC transporter required for expression of cytochrome bd (ATP-) cydD - Bacillus subtilis dbj BAA11730.1 (D83026) homologous to many ATP-binding transport proteins; hypothetical [Bacillus subtilis] emb CAB15899.1 (Z99123) ABC membrane transporter (ATP-binding protein) [Bacillus subtilis] Length = 575</p>
SeqID 440	SA-1313.1	Contig128 (33946-35664 m)	67	<p>Identities = 278/569 (48%), Positives = 399/569 (69%), Gaps = 6/569 (1%) sp P94366 CYDC_BACSU TRANSPORT ATP-BINDING PROTEIN CYDC pir C69611 ABC transporter required for expression of cytochrome bd (ATP-) cydC - Bacillus subtilis dbj BAA11729.1 (D83026) homologous to many ATP-binding transport proteins including binding transport proteins including SwissProt:CYDD_ECOLI; hypothetical [Bacillus subtilis] emb CAB15900.1 (Z99123) ABC membrane transporter (ATP-binding protein) [Bacillus subtilis] Length = 567</p>

SeqID 441	SA-1314.1	Contig128 (35664-36683 m)	65	<p>Identities = 158/331 (47%), Positives = 223/331 (66%), Gaps = 1/331 (0%) sp P94365 CYDB_BACSU CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT II pir B69811 cytochrome d ubiquinol oxidase (EC 1.10.3.-) chain II cydB - Bacillus subtilis dbj BAA11728.1 (D83026) homologous to cytochrome d ubiquinol oxidase subunit II; hypothetical [Bacillus subtilis] emb CAB15901.1 (Z99123) cytochrome bd ubiquinol oxidase (subunit II) [Bacillus subtilis] Length = 338</p>
SeqID 442	SA-1316.1	Contig128 (36684-38111 m)	65	<p>Identities = 246/470 (52%), Positives = 319/470 (67%), Gaps = 12/470 (2%) sp P94364 CYDA_BACSU CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT I pir A69611 cytochrome bd ubiquinol oxidase (subunit I) cydA - Bacillus subtilis dbj BAA11727.1 (D83026) homologous to cytochrome d ubiquinol oxidase subunit I; hypothetical [Bacillus subtilis] emb CAB15902.1 (Z99123) cytochrome bd ubiquinol oxidase (subunit I) [Bacillus subtilis] Length = 468</p>
SeqID 443	SA-1318.1	Contig128 (38214-39422 m)	61	<p>Identities = 178/403 (44%), Positives = 249/403 (61%), Gaps = 7/403 (1%) pir A70015 probable NADH dehydrogenase (EC 1.6.99.3) yumB - Bacillus subtilis emb CAB07953.1 (Z93939) unknown [Bacillus subtilis] emb CAB15200.1 (Z99120) similar to NADH dehydrogenase [Bacillus subtilis] Length = 406</p>

SeqID 444	SA-1319.2	Contig128 (39435-40334 m)	44	Identities = 74/290 (25%), Positives = 138/290 (47%), Gaps = 15/290 (5%) sp P39582 MENA_BACSU PROBABLE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (DHNA-OCTAPRENYLTRANSFERASE) pir S39661 menaquinone biosynthesis protein homolog ywaB - Bacillus subtilis emb CAA51562.1 (X73124) ipa-6d [Bacillus subtilis] emb CAB15875.1 (Z99123) alternate gene name: ipa-6d~similar to quinone biosynthesis [Bacillus subtilis] Length = 311
SeqID 445	SA-132.2	Contig138 (94727-95020 m)	42	Identities = 35/72 (48%), Positives = 42/72 (57%) pir S59084 hypothetical protein 29.1 - red alga (Chondrus crispus) mitochondrion emb CAA87600.1 (Z47547) unique orf [Chondrus crispus] Length = 79
SeqID 446	SA-1320.1	Contig100 (14165-14662 p)	No Hits found	
SeqID 447	SA-1322.1	Contig100 (13666-14115 p)	No Hits found	
SeqID 448	SA-1323.1	Contig100 (13349-13642 p)	No Hits found	
SeqID 449	SA-1324.1	Contig100 (12954-13271 p)	67	Identities = 70/96 (72%), Positives = 83/96 (85%) gb AAB52379.1 (U31811) immunogenic secreted protein precursor [Streptococcus pyogenes] Length = 534
SeqID 450	SA-1325.1	Contig100 (12107-13018 p)	60	Identities = 142/298 (47%), Positives = 185/298 (61%), Gaps = 11/298 (3%) gb AAB52379.1 (U31811) immunogenic secreted protein precursor [Streptococcus pyogenes] Length = 534
SeqID 451	SA-1326.1	Contig100 (11805-12203 p)	32	Identities = 35/112 (31%), Positives = 49/112 (43%), Gaps = 12/112 (10%) pir T07945 probable arabinogalactan protein (clone Sta 39-3) - rape gb AAC37509.1 (L47351) arabinogalactan protein [Brassica napus] Length = 136
SeqID 452	SA-1327.1	Contig100 (11394-11648 p)	57	Identities = 32/76 (42%), Positives = 54/76 (70%) dbj BAB04699.1 (AP001510) unknown conserved protein [Bacillus halodurans] Length = 102

SeqID 453	SA-1328.1	Contig100 (10783-11373 p)	55	Identities = 72/259 (27%), Positives = 112/259 (42%), Gaps = 64/259 (24%) pir C82882 ABC Transporter UU510 [imported] - Ureaplasma urealyticum gb AAF30922.1 AE002149_7 (AE002149) ABC Transporter [Ureaplasma urealyticum] Length = 658
SeqID 454	SA-1329.1	Contig100 (10314-10769 p)	No Hits found	
SeqID 455	SA-133.1	Contig138 (95173-95493 p)	63	Identities = 62/104 (59%), Positives = 68/104 (64%) pir F81516 hypothetical protein CP0987 [imported] - Chlamydomonas pneumoniae (strain AR39) Length = 216
SeqID 456	SA-1330.1	Contig100 (9421-10335 p)	34	Identities = 59/254 (23%), Positives = 108/254 (42%), Gaps = 14/254 (5%) emb CAC16670.1 (AJ302698) hypothetical protein [Staphylococcus haemolyticus] Length = 373
SeqID 457	SA-1331.1	Contig100 (9062-9418 p)	No Hits found	
SeqID 458	SA-1332.1	Contig100 (8700-9050 p)	No Hits found	
SeqID 459	SA-1333.2	Contig100 (8666-8776 m)	No Hits found	
SeqID 460	SA-1334.1	Contig100 (4757-8686 p)	50	Identities = 452/1058 (42%), Positives = 664/1058 (62%), Gaps = 39/1058 (3%) pir C70013 conserved hypothetical protein yuka - Bacillus subtilis emb CAB15175.1 (Z99120) alternate gene name: yueA~similar to hypothetical proteins [Bacillus subtilis] Length = 1207
SeqID 461	SA-1335.1	Contig100 (4279-4782 p)	No Hits found	
SeqID 462	SA-1336.1	Contig100 (2998-4272 p)	37	Identities = 83/323 (25%), Positives = 163/323 (49%), Gaps = 1/323 (0%) dbj BAB04693.1 (AP001510) unknown conserved protein [Bacillus halodurans] Length = 440
SeqID 463	SA-1337.1	Contig100 (2756-2998 p)	No Hits found	
SeqID 464	SA-1338.1	Contig100 (2344-2772 p)	No Hits found	
SeqID 465	SA-1339.1	Contig100 (2-2290 p)	27	Identities = 101/501 (20%), Positives = 213/501 (42%), Gaps = 54/501 (10%) pir PC6003 surface membrane protein Imp4 - Mycoplasma hominis (fragment) Length = 624
SeqID 466	SA-134.1	Contig138 (94635-94964 m)	No Hits found	

SeqID 467	SA-1340.2	Contig120 (16456-17268 m)	58	Identities = 116/260 (44%), Positives = 165/260 (62%), Gaps = 6/260 (2%) sp P39345 IDNO_ECOLI_GLUCONATE 5-DEHYDROGENASE (5-KETO-D-GLUCONATE 5-REDUCTASE) pir S56492 hypothetical 27.6K protein (pepa-gntv intergenic region) - Escherichia coli gb AAA97163.1 (U14003) ORF_1254 [Escherichia coli] gb AAC77223.1 (AE000497) 5-keto-D-gluconate 5-reductase [Escherichia coli K12] Length = 254
SeqID 468	SA-1341.1	Contig120 (15801-16439 m)	38	Identities = 47/189 (24%), Positives = 81/189 (41%), Gaps = 20/189 (10%) pir D43258 galactose-6-phosphate isomerase subunit LacB - Streptococcus mutans Length = 171
SeqID 469	SA-1342.1	Contig120 (14768-15775 m)	48	Identities = 94/329 (28%), Positives = 169/329 (50%), Gaps = 7/329 (2%) pir G72422 2-keto-3-deoxygluconate kinase - Thermotoga maritima (strain MSB8) gb AAD35161.1 AE001693_7 (AE001693) 2-keto-3-deoxygluconate kinase [Thermotoga maritima] Length = 339
SeqID 470	SA-1343.1	Contig120 (14118-14756 m)	51	Identities = 78/192 (40%), Positives = 118/192 (60%), Gaps = 6/192 (3%) pir F72422 KHG-KDPG bifunctional aldolase TM0066 [similarity] - Thermotoga maritima (strain MSB8) gb AAD35160.1 AE001693_6 (AE001693) 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase [Thermotoga maritima] Length = 205
SeqID 471	SA-1344.2	Contig120 (11557-13410 m)	35	Identities = 112/529 (21%), Positives = 218/529 (41%), Gaps = 67/529 (12%) emb CAB62846.2 (AL035475) hypothetical protein [Plasmodium falciparum] Length = 1734
SeqID 472	SA-1345.1	Contig120 (10772-11386 p)	67	Identities = 115/203 (56%), Positives = 143/203 (69%), Gaps = 7/203 (3%) pir A83323 hypothetical protein PA2575 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG05963.1 AE004686_1 (AE004686) hypothetical protein [Pseudomonas aeruginosa] Length = 200

SeqID 473	SA-1347.1	Contig120 (10214-10639 p)	48	Identities = 49/124 (39%), Positives = 73/124 (58%) pir A69860 transcription regulator MarR family homolog ykoM - Bacillus subtilis emb CAA05611.1 (AJ002571) YkoM [Bacillus subtilis] emb CAB13191.1 (Z99110) similar to transcriptional regulator (MarR family) [Bacillus subtilis] emb CAB13207.1 (Z99111) similar to transcriptional regulator (MarR family) [Bacillus subtilis] Length = 154
SeqID 474	SA-1348.1	Contig120 (5681-10087 p)	91	Identities = 1238/1468 (84%), Positives = 1346/1468 (91%), Gaps = 3/1468 (0%) gb AAF98345.1 (AF280761) DNA polymerase III alpha chain [Streptococcus pyogenes] Length = 1465
SeqID 475	SA-1349.1	Contig120 (4974-5558 p)	61	Identities = 101/201 (50%), Positives = 122/201 (60%), Gaps = 9/201 (4%) emb CAB94815.1 (AJ245582) peptidoglycan hydrolase [Streptococcus thermophilus] Length = 218
SeqID 476	SA-135.1	Contig138 (93460-93819 p)	No Hits found	
SeqID 477	SA-1350.1	Contig120 (2996-4849 p)	65	Identities = 300/608 (49%), Positives = 410/608 (67%), Gaps = 52/608 (8%) sp O31755 SYP_BACSU PROLYL-TRNA SYNTHETASE (PROLINE--TRNA LIGASE) (PRORS) pir G69682 proline--tRNA ligase (EC 6.1.1.15) proS - Bacillus subtilis emb CAB13530.1 (Z99112) prolyl-tRNA synthetase [Bacillus subtilis] Length = 564
SeqID 478	SA-1351.1	Contig120 (1645-2904 p)	70	Identities = 229/425 (53%), Positives = 298/425 (69%), Gaps = 9/425 (2%) gb AAD47948.1 AF152237_1 (AF152237) Eep [Enterococcus faecalis] Length = 422
SeqID 479	SA-1352.1	Contig120 (820-1614 p)	63	Identities = 116/266 (43%), Positives = 172/266 (64%), Gaps = 6/266 (2%) dbj BAB06141.1 (AP001515) phosphatidate cytidyltransferase [Bacillus halodurans] Length = 264

SeqID 480	SA-1353.1	Contig120 (53-805 p)	63	Identities = 134/235 (57%), Positives = 170/235 (72%), Gaps = 4/235 (1%) sp O31751 UPPS_BACSU UNDECAPRENYL PYROPHOSPHATE SYNTHETASE (UPP SYNTHETASE) (DI-TRANS-POLY-CIS-DECAPRENYLICISTRANSFERASE) (UNDECAPRENYL DIPHOSPHATE SYNTHASE) (UDS) pir A69881 conserved hypothetical protein yluA - Bacillus subtilis emb CAB13526.1 (Z99112) similar to hypothetical proteins [Bacillus subtilis] Length = 260
SeqID 481	SA-1354.1	Contig99 (17810-18058 p)	No Hits found	
SeqID 482	SA-1355.1	Contig99 (16154-17593 p)	76	Identities = 289/478 (60%), Positives = 368/478 (76%), Gaps = 2/478 (0%) gb AAA69226.1 (U29579) 6-phospho-beta-glucosidase [Escherichia coli] Length = 476
SeqID 483	SA-1356.1	Contig99 (15380-15994 p)	40	Identities = 43/177 (24%), Positives = 84/177 (47%), Gaps = 9/177 (5%) sp O35264 PA1B_RAT PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE IB BETA SUBUNIT (PAF ACETYLHYDROLASE 30 KDA SUBUNIT) (PAF-AH 30 KDA SUBUNIT) (PAF-AH BETA SUBUNIT) (PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE ALPHA 2 SUBUNIT) (PAF-AH ALPHA 2) gb AAC27974.1 (AF016048) platelet-activating factor acetylhydrolase alpha 2 subunit [Rattus norvegicus] Length = 229
SeqID 484	SA-1357.1	Contig99 (14458-15279 p)	55	Identities = 103/265 (38%), Positives = 154/265 (57%), Gaps = 4/265 (1%) sp P75809 YBJI_ECOLI PROTEIN YBJI Length = 271
SeqID 485	SA-1358.1	Contig99 (12904-13836 m)	75	Identities = 197/296 (66%), Positives = 240/296 (80%) gb AAF89977.1 AF206272_3 (AF206272) transcriptional regulator [Streptococcus mutans] Length = 301
SeqID 486	SA-1359.1	Contig99 (11315-12847 p)	46	Identities = 141/443 (31%), Positives = 241/443 (53%), Gaps = 20/443 (4%) pir C82449 conserved hypothetical protein VCA0526 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF96429.1 (AE004383) conserved hypothetical protein [Vibrio cholerae] Length = 468

SeqID 487	SA-136.1	Contig138 (92309-93091 m)	No Hits found	
SeqID 488	SA-1361.1	Contig99 (10178-11206 p)	50	<p>Identities = 131/336 (38%), Positives = 188/336 (54%), Gaps = 7/336 (2%) sp P23861 POTD_ECOLI SPERMIDINE/PUTRESCINE-BINDING PERIPLASMIC PROTEIN PRECURSOR (SPBP) pir D40840 spermidine/putrescine-binding protein precursor [validated] - Escherichia coli gb AAC37041.1 (M64519) transport protein [Escherichia coli] dbj BAA35943.1 (D90747) Spermidine/putrescine transport protein D [Escherichia coli] gb AAC74207.1 (AE000212) spermidine/putrescine periplasmic transport protein [Escherichia coli K12] Length = 348</p>
SeqID 489	SA-1362.1	Contig99 (9364-10140 p)	55	<p>Identities = 97/249 (38%), Positives = 159/249 (62%), Gaps = 3/249 (1%) pir G70179 spermidine/putrescine ABC transporter, permease protein (potC) homolog - Lyme disease spirochete gb AAB91527.1 (AE001165) spermidine/putrescine ABC transporter, permease protein (potC) [Borrelia burgdorferi] Length = 263</p>
SeqID 490	SA-1363.1	Contig99 (8573-9367 p)	56	<p>Identities = 90/255 (35%), Positives = 153/255 (59%), Gaps = 11/255 (4%) sp P45170 POTB_HAEIN SPERMIDINE/PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POTB pir A64118 spermidine/putrescine transport system permease potB - Haemophilus influenzae (strain Rd KW20) gb AAC22990.1 (U32813) spermidine/putrescine ABC transporter, permease protein (potB) [Haemophilus influenzae Rd] Length = 286</p>
SeqID 491	SA-1364.1	Contig99 (7435-8589 p)	62	<p>Identities = 165/345 (47%), Positives = 240/345 (68%), Gaps = 1/345 (0%) pir A70180 spermidine/putrescine ABC transporter, ATP-binding protein (potA) homolog - Lyme disease spirochete gb AAB91525.1 (AE001165) spermidine/putrescine ABC transporter, ATP-binding protein (potA) [Borrelia burgdorferi] Length = 347</p>

SeqID 492	SA-1365.1	Contig99 (6484-7386 p)	53	Identities = 119/286 (41%), Positives = 166/286 (57%), Gaps = 1/286 (0%) dbj BAB06283.1 (AP001515) UDP-N-acetylpyruvoylglucosamine reductase [Bacillus halodurans] Length = 301
SeqID 493	SA-1366.1	Contig99 (5852-6340 p)	50	Identities = 65/131 (49%), Positives = 86/131 (65%) sp P29252 HPPK_BACSU 2-AMINO-4-HYDROXY-6-HYDROXYMETHYLDIHYDROPTERIDINE PYROPHOSPHOKINASE (7,8-DIHYDRO-6-HYDROXYMETHYLPTERIN-PYROPHOSPHOKINASE) (HPPK) (6-HYDROXYMETHYL-7,8-DIHYDROPTERIN PYROPHOSPHOKINASE) (PPP) pir S66109 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (EC 2.7.6.3) - Bacillus subtilis dbj BAA05314.1 (D26185) 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase [Bacillus subtilis] emb CAB11855.1 (Z99104) 7,8-dihydro-6-hydroxymethylpterin pyrophosphokinase [Bacillus subtilis] Length = 167
SeqID 494	SA-1367.1	Contig99 (5493-5855 p)	70	Identities = 72/119 (60%), Positives = 90/119 (75%) sp O33725 FOLB_STRPY DIHYDRONEOPTERIN ALDOLASE (DHNA) emb CAA04239.1 (AJ000685) dihydroneopterin aldolase [Streptococcus pyogenes] Length = 119
SeqID 495	SA-1368.1	Contig99 (4688-5491 p)	83	Identities = 182/267 (68%), Positives = 224/267 (83%), Gaps = 1/267 (0%) emb CAA04242.1 (AJ000686) dihydropteroate synthase [Streptococcus pyogenes] Length = 266
SeqID 496	SA-137.1	Contig138 (91179-92312 p)	59	Identities = 154/382 (40%), Positives = 224/382 (58%), Gaps = 19/382 (4%) pir A69774 integrase homolog ydcL - Bacillus subtilis dbj BAA19318.1 (AB001488) PROBABLE INTEGRASE. [Bacillus subtilis] emb CAB12287.1 (Z99106) similar to integrase [Bacillus subtilis] Length = 368

SeqID 497	SA-1370.1	Contig99 (4121-4684 p)	87	<p>Identities = 148/184 (80%), Positives = 168/184 (90%) sp O33723 GCH1_STRPY GTP CYCLOHYDROLASE I (GTP-CH- I) emb CAA04237.1 (AJ00685) GTP cyclohydrolase [Streptococcus pyogenes] Length = 194</p>
SeqID 498	SA-1371.1	Contig99 (2840-4102 p)	57	<p>Identities = 154/426 (36%), Positives = 245/426 (57%), Gaps = 17/426 (3%) sp Q05865 FOLC_BACSU FOLYLPOLYGLUTAMATE SYNTHASE (FOLYLPOLY-GAMMA- GLUTAMATE SYNTHETASE) (FPGS) pir B40646 folyl- polyglutamate synthetase folC - Bacillus subtilis gb AAB5902.1 (L04520) folyl-polyglutamate synthetase [Bacillus subtilis] emb CAB14768.1 (Z99118) folyl-polyglutamate synthetase [Bacillus subtilis] Length = 430</p>
SeqID 499	SA-1372.1	Contig99 (1951-2838 p)	63	<p>Identities = 110/295 (37%), Positives = 188/295 (63%), Gaps = 6/295 (2%) dbj BAB07585.1 (AP001520) unknown conserved protein [Bacillus halodurans] Length = 308</p>
SeqID 500	SA-1373.2	Contig99 (1098-1964 p)	79	<p>Identities = 188/289 (65%), Positives = 232/289 (80%), Gaps = 1/289 (0%) sp P72535 KHSE_STRPN HOMOSERINE KINASE (HK) gb AAC44297.1 (U41735) homoserine kinase homolog [Streptococcus pneumoniae] Length = 289</p>
SeqID 501	SA-1374.2	Contig99 (2-1096 p)	71	<p>Identities = 187/368 (50%), Positives = 261/368 (70%), Gaps = 11/368 (2%) sp P52985 DHOM_LACLA HOMOSERINE DEHYDROGENASE (HDH) pir JC6049 homoserine dehydrogenase (EC 1.1.1.3) - Lactococcus lactis emb CAA65713.1 (X96988) hom [Lactococcus lactis] Length = 428</p>
SeqID 502	SA-1377.1	Contig134 (51276-52901 p)	59	<p>Identities = 210/541 (38%), Positives = 326/541 (59%), Gaps = 14/541 (2%) dbj BAB04286.1 (AP001509) nickel transport system (nickel-binding protein) [Bacillus halodurans] Length = 539</p>

SeqID 503	SA-1378.1	Contig134 (52888-53832 p)	55	Identities = 121/304 (39%), Positives = 176/304 (57%) dbj BAB04287.1 (AP001509) nickel transport system (permease) [Bacillus halodurans] Length = 314
SeqID 504	SA-1379.1	Contig134 (53871-54638 p)	58	Identities = 106/255 (41%), Positives = 164/255 (63%) dbj BAB04288.1 (AP001509) nickel transport system (permease) [Bacillus halodurans] Length = 277
SeqID 505	SA-138.1	Contig138 (90969-91175 p)	No Hits found	
SeqID 506	SA-1380.1	Contig134 (54614-55414 p)	57	Identities = 85/253 (33%), Positives = 154/253 (60%), Gaps = 2/253 (0%) gb AAF73561.1 (AE002315) peptide ABC transporter, ATP-binding protein [Chlamydia muridarum] Length = 281
SeqID 507	SA-1381.1	Contig134 (55401-56081 p)	55	Identities = 81/199 (40%), Positives = 131/199 (65%), Gaps = 2/199 (1%) dbj BAB05797.1 (AP001514) oligopeptide ABC transporter (ATP-binding protein) [Bacillus halodurans] Length = 244
SeqID 508	SA-1382.1	Contig134 (56202-56930 p)	76	Identities = 143/238 (60%), Positives = 193/238 (81%) sp O31749 PYRH_BACSU URIDYLATE KINASE (UK) (URIDINE MONOPHOSPHATE KINASE) (UMP KINASE) pir F69708 uridylylate kinase smbA - Bacillus subtilis emb CAB13524.1 (Z99112) uridylylate kinase [Bacillus subtilis] Length = 240
SeqID 509	SA-1383.1	Contig134 (56946-57503 p)	78	Identities = 112/185 (60%), Positives = 149/185 (80%) dbj BAB06143.1 (AP001515) ribosome recycling factor [Bacillus halodurans] Length = 185
SeqID 510	SA-1384.1	Contig134 (57621-58475 p)	49	Identities = 107/269 (39%), Positives = 155/269 (56%), Gaps = 6/269 (2%) pir E69840 hypothetical protein yitL - Bacillus subtilis emb CAB12943.1 (Z99109) yitL [Bacillus subtilis] Length = 298
SeqID 511	SA-1385.1	Contig134 (58601-59122 p)	67	Identities = 102/175 (58%), Positives = 127/175 (72%), Gaps = 2/175 (1%) dbj BAB05167.1 (AP001512) peptide methionine sulfoxide reductase [Bacillus halodurans] Length = 179

SeqID 512	SA-1386.1	Contig134 (59143-59334 p)	50	Identities = 24/66 (36%), Positives = 42/66 (63%) pir A69931 hypothetical protein yozE - Bacillus subtilis emb CAB13859.1 (Z99114) yozE [Bacillus subtilis] Length = 74
SeqID 513	SA-1388.1	Contig134 (59492-61264 p)	89	Identities = 472/590 (80%), Positives = 535/590 (90%) gb AAA51870.1 (U09352) 67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus pyogenes] Length = 590
SeqID 514	SA-1389.1	Contig134 (61388-62374 p)	65	Identities = 191/305 (62%), Positives = 241/305 (78%), Gaps = 1/305 (0%) sp P46343 PHOL_BACSU PHOH-LIKE PROTEIN pir E69676 phosphate starvation-induced protein phoH - Bacillus subtilis dbj BAA12477.1 (D84432) YqfE [Bacillus subtilis] emb CAB14476.1 (Z99117) phosphate starvation-induced protein [Bacillus subtilis] Length = 319
SeqID 515	SA-1390.2	Contig134 (62465-63268 m)	25	Identities = 42/161 (26%), Positives = 72/161 (44%), Gaps = 19/161 (11%) gb AAC95491.1 (U73025) unknown [Staphylococcus aureus] gb AAG42230.1 AF299292_4 (AF299292) ORFX [Staphylococcus intermedius] Length = 289
SeqID 516	SA-1391.2	Contig133 (46310-47965 m)	56	Identities = 176/589 (29%), Positives = 315/589 (52%), Gaps = 47/589 (7%) dbj BAB06875.1 (AP001517) two-component sensor histidine kinase involved in phosphate regulation [Bacillus halodurans] Length = 589
SeqID 517	SA-1392.1	Contig133 (47958-48635 m)	59	Identities = 98/224 (43%), Positives = 138/224 (60%), Gaps = 2/224 (0%) sp P45606 PHOB_SHIDY PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN PHOB pir A44753 phoB protein - Shigella dysenteriae gb AAA26535.1 (M31793) phosphate regulatory protein phoB [Shigella dysenteriae] Length = 229

SeqID 518	SA-1393.1	Contig133 (48635-49291 m)	51	Identities = 66/213 (30%), Positives = 119/213 (54%), Gaps = 4/213 (1%) sp Q51547 PHOU_PSEAE PHOSPHATE TRANSPORT SYSTEM PROTEIN PHOU pir S65576 negative regulator PhoU - Pseudomonas aeruginosa pir S68596 negative regulator PhoU - Pseudomonas aeruginosa pir H82975 phosphate uptake regulatory protein PhoU PA5365 [imported] - Pseudomonas aeruginosa (strain PAO1) dbj BAA08138.1 (D45195) a negative regulator of pho regulon [Pseudomonas aeruginosa] gb AAG08750.1 AE004948_6 (AE004948) phosphate uptake regulatory protein PhoU [Pseudomonas aeruginosa] Length = 242
SeqID 519	SA-1394.1	Contig133 (49288-50037 m)	82	Identities = 166/245 (67%), Positives = 211/245 (85%), Gaps = 1/245 (0%) gb AAD22041.1 (AF118229) ATP-binding cassette protein PstB [Streptococcus pneumoniae] Length = 250
SeqID 520	SA-1395.1	Contig133 (50030-50908 m)	66	Identities = 135/263 (51%), Positives = 203/263 (76%) gb AAD22040.1 (AF118229) transmembrane protein PstA [Streptococcus pneumoniae] Length = 271
SeqID 521	SA-1396.1	Contig133 (50910-51755 m)	74	Identities = 162/266 (60%), Positives = 212/266 (78%), Gaps = 3/266 (1%) gb AAD22039.1 (AF118229) transmembrane protein PstC [Streptococcus pneumoniae] Length = 271
SeqID 522	SA-1398.2	Contig133 (52368-52664 p)	No Hits found	
SeqID 523	SA-1399.1	Contig133 (51770-52651 m)	77	Identities = 230/230 (100%), Positives = 230/230 (100%) pir A61607 probable hemolysin precursor - Streptococcus agalactiae (strain 74-360) Length = 230
SeqID 524	SA-14.1	Contig137 (32042-32473 p)	32	Identities = 31/118 (26%), Positives = 51/118 (42%), Gaps = 3/118 (2%) pir S67490 single-stranded DNA-binding protein - Eubacterium sp gb AAA79866.1 (U12515) single-stranded DNA binding protein [uncultured eubacterium] prf J2108276A ssDNA-binding protein [Rattus norvegicus] Length = 181

SeqID 525	SA-140.1	Contig138 (90097-90963 p)	47	Identities = 86/278 (30%), Positives = 137/278 (48%), Gaps = 36/278 (12%) dbj BAA07788.1 (D43692) rep protein [Brevibacillus borstelensis] Length = 281
SeqID 526	SA-1400.1	Contig133 (52852-53439 m)	No Hits found	
SeqID 527	SA-1401.1	Contig133 (53436-54176 m)	58	Identities = 119/250 (47%), Positives = 149/250 (59%), Gaps = 9/250 (3%) dbj BAB05069.1 (AP001511) unknown conserved protein [Bacillus halodurans] Length = 250
SeqID 528	SA-1403.1	Contig133 (54176-55129 m)	59	Identities = 138/309 (44%), Positives = 193/309 (61%), Gaps = 5/309 (1%) pir T43740 probable ribosomal protein L11 methyltransferase (EC 2.1.1.-) [imported] - Listeria monocytogenes dbj BAA82791.1 (AB023064) orf35 [Listeria monocytogenes] Length = 314
SeqID 529	SA-1404.1	Contig133 (55126-55434 m)	No Hits found	
SeqID 530	SA-1405.1	Contig133 (55708-56424 p)	57	Identities = 87/246 (35%), Positives = 140/246 (56%), Gaps = 13/246 (5%) pir C69661 transcription activator of multidrug-efflux transporter genes mta - Bacillus subtilis emb CAA69863.1 (Y08559) Unknown [Bacillus subtilis] emb CAB15677.1 (Z99122) transcriptional regulator [Bacillus subtilis] Length = 257
SeqID 531	SA-1406.1	Contig133 (56463-56933 m)	36	Identities = 38/136 (27%), Positives = 60/136 (43%), Gaps = 3/136 (2%) sp P54441 YRKN_BACSU HYPOTHETICAL 21.3 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION pir D69977 hypothetical protein yrkN - Bacillus subtilis dbj BAA12369.1 (D84432) YrkN [Bacillus subtilis] emb CAB14586.1 (Z99117) yrkN [Bacillus subtilis] Length = 185
SeqID 532	SA-1407.1	Contig133 (56905-57363 m)	52	Identities = 57/131 (43%), Positives = 82/131 (62%) pir F83247 hypothetical protein PA3180 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG06568.1 AE004742_4 (AE004742) hypothetical protein [Pseudomonas aeruginosa] Length = 145
SeqID 533	SA-1408.1	Contig133 (57350-57496 m)	No Hits found	
SeqID 534	SA-141.1	Contig138 (89680-89994 p)	No Hits found	

SeqID 535	SA-1410.1	Contig133 (57499-57969 m)	No Hits found	
SeqID 536	SA-1413.1	Contig97 (10056-12434 p)	61	Identities = 360/785 (45%), Positives = 496/785 (62%), Gaps = 15/785 (1%) pir S76896 hypothetical protein - <i>Synechocystis</i> sp. (strain PCC 6803) dbj BAA18808.1 (D90917) hypothetical protein [Synechocystis sp.] Length = 821
SeqID 537	SA-1414.1	Contig97 (8875-9966 p)	73	Identities = 211/363 (58%), Positives = 270/363 (74%), Gaps = 9/363 (2%) sp P39300 YJFR_ECOLI_HYPOTHETICAL 40.1 KD PROTEIN IN AIDB-SGAT INTERGENIC REGION Length = 354
SeqID 538	SA-1415.1	Contig97 (6879-8558 p)	41	Identities = 136/511 (26%), Positives = 234/511 (45%), Gaps = 29/511 (5%) dbj BAB03939.1 (AP001507) unknown conserved protein [Bacillus halodurans] Length = 701
SeqID 539	SA-1416.1	Contig97 (6052-6813 p)	39	Identities = 58/235 (24%), Positives = 104/235 (43%), Gaps = 9/235 (3%) pir C83362 hypothetical protein PA2260 [imported] - <i>Pseudomonas aeruginosa</i> (strain PAO1) gb AAG05648.1 AE004652_1 (AE004652) hypothetical protein [Pseudomonas aeruginosa] Length = 260
SeqID 540	SA-1417.1	Contig97 (5766-6032 p)	40	Identities = 28/89 (31%), Positives = 44/89 (48%), Gaps = 3/89 (3%) sp P44991 LYXK_HAEIN PROBABLE L-XYLULOSE KINASE (L-XYLULOSE KINASE) pir H64164 hypothetical protein HI1027 - <i>Haemophilus influenzae</i> (strain Rd KW20) gb AAC22687.1 (U32783) L-xylulose kinase (lyx) [Haemophilus influenzae Rd] Length = 485
SeqID 541	SA-1418.1	Contig97 (4513-5727 p)	48	Identities = 126/393 (32%), Positives = 212/393 (53%), Gaps = 16/393 (4%) sp P37677 LYXK_ECOLI_CRYPTIC L-XYLULOSE KINASE (L-XYLULOSE KINASE) pir S47801 L-xylulokinase [EC 2.7.1.53] - <i>Escherichia coli</i> gb AAB18557.1 (U00039) No definition line found [Escherichia coli] gb AAC76604.1 (AE000435) L-xylulose kinase, cryptic [Escherichia coli K12] Length = 498
SeqID 542	SA-1419.1	Contig97 (3458-4447 p)	No Hits found	

SeqID 543	SA-1421.1	Contig97 (1985-3436 p)	40	Identities = 106/402 (26%), Positives = 196/402 (48%), Gaps = 15/402 (3%) dbj BAA75340.1 (AB011837) PTS system galactitol-specific enzyme IIC component [Bacillus halodurans] dbj BAB03909.1 (AP001507) PTS system, galactitol-specific enzyme II, C component [Bacillus halodurans] Length = 419
SeqID 544	SA-1424.1	Contig97 (931-1887 p)	59	Identities = 123/326 (37%), Positives = 193/326 (58%), Gaps = 8/326 (2%) pir B75057 glycerate dehydrogenase PAB2374 - Pyrococcus abyssi (strain Orsay) emb CAB50351.1 (AJ248287) GLYCERATE DEHYDROGENASE [Pyrococcus abyssi] Length = 335
SeqID 545	SA-1425.1	Contig97 (1-912 p)	No Hits found	
SeqID 546	SA-1427.1	Contig115 (8540-9379 p)	57	Identities = 101/275 (36%), Positives = 165/275 (59%), Gaps = 22/275 (8%) gb AAD50462.1 AF169967_5 (AF169967) BacA [Flavobacterium johnsoniae] Length = 265
SeqID 547	SA-1429.1	Contig115 (9499-10254 p)	80	Identities = 168/248 (67%), Positives = 205/248 (81%), Gaps = 9/248 (3%) dbj BAA82113.1 (AB022909) negative regulator of genetic competence [Streptococcus mutans] Length = 240
SeqID 548	SA-143.1	Contig138 (88238-89581 p)	40	Identities = 103/460 (22%), Positives = 191/460 (41%), Gaps = 72/460 (15%) pir F69774 hypothetical protein ydcQ [imported] - Bacillus subtilis dbj BAA19323.1 (AB001488) SIMILAR TO ORF21 OF ENTEROCOCCUS FAECALIS TRANSPOSON TN916. [Bacillus subtilis] emb CAB12293.1 (Z99106) similar to transposon protein [Bacillus subtilis] Length = 480
SeqID 549	SA-1430.1	Contig115 (10400-11416 p)	80	Identities = 267/382 (69%), Positives = 317/382 (82%) dbj BAA82114.1 (AB022909) RgpG [Streptococcus mutans] Length = 388

SeqID 550	SA-1431.1	Contig115 (11581-12351 p)	77	Identities = 180/250 (72%), Positives = 212/250 (84%) sp P80866 V296_BACSU VEGETATIVE PROTEIN 296 (VEG296) pir H70019 ABC transporter (ATP-binding protein) homolog yurY - Bacillus subtilis emb CAB15260.1 (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 261
SeqID 551	SA-1432.1	Contig115 (12388-13650 p)	62	Identities = 174/435 (40%), Positives = 269/435 (61%), Gaps = 15/435 (3%) dbj BAB07189.1 (AP001518) unknown conserved protein [Bacillus halodurans] Length = 435
SeqID 552	SA-1433.1	Contig115 (13652-14884 p)	74	Identities = 241/400 (60%), Positives = 308/400 (76%), Gaps = 5/400 (1%) pir F70019 nifS protein homolog yurW - Bacillus subtilis emb CAB15258.1 (Z99120) similar to NifS protein homolog [Bacillus subtilis] Length = 406
SeqID 553	SA-1434.1	Contig115 (14871-15314 p)	58	Identities = 72/139 (51%), Positives = 92/139 (65%) sp O32163 NIFU_BACSU NIFU-LIKE PROTEIN pir E70019 nitrogen fixation protein nifU homolog yurV - Bacillus subtilis emb CAB15257.1 (Z99120) similar to NifU protein homolog [Bacillus subtilis] Length = 147
SeqID 554	SA-1435.1	Contig115 (15414-16832 p)	79	Identities = 315/459 (68%), Positives = 385/459 (83%) pir D70019 conserved hypothetical protein yurU - Bacillus subtilis emb CAB15256.1 (Z99120) similar to hypothetical proteins [Bacillus subtilis] Length = 465
SeqID 555	SA-1436.1	Contig115 (16904-18031 m)	45	Identities = 117/333 (35%), Positives = 187/333 (56%), Gaps = 8/333 (2%) emb CAA60585.1 (X87105) penicillin binding protein 4 [Staphylococcus aureus] emb CAA60582.1 (X87104) penicillin binding protein 4 [Staphylococcus aureus] Length = 431
SeqID 556	SA-1437.2	Contig115 (18244-19479 m)	71	Identities = 203/415 (48%), Positives = 300/415 (71%), Gaps = 6/415 (1%) emb CAA67776.1 (X99400) D,D-carboxypeptidase [Streptococcus pneumoniae] Length = 413

SeqID 557	SA-1438.2	Contig139 (86184-86444 m)	71	Identities = 61/81 (75%), Positives = 71/81 (87%), Gaps = 1/81 (1%) gb AAF80389.1 AF160251_3 (AF160251) ribosomal protein L31 [Listeria innocua] Length = 81
SeqID 558	SA-1439.1	Contig139 (84551-86071 p)	61	Identities = 257/429 (59%), Positives = 312/429 (71%), Gaps = 7/429 (1%) pir T46756 Zn-binding lipoprotein adcA [imported] - Streptococcus pneumoniae (fragment) emb CAA96185.1 (Z71552) AdcA protein [Streptococcus pneumoniae] Length = 423
SeqID 559	SA-144.1	Contig138 (87835-88260 p)	No Hits found	
SeqID 560	SA-1440.1	Contig139 (83012-84409 p)	65	Identities = 218/473 (46%), Positives = 310/473 (65%), Gaps = 14/473 (2%) emb CAA86210.1 (Z38063) dipeptidase [Lactobacillus helveticus] Length = 474
SeqID 561	SA-1441.1	Contig139 (82086-82997 p)	82	Identities = 209/303 (68%), Positives = 260/303 (84%) emb CAB96619.1 (AJ400630) hypothetical protein [Streptococcus pneumoniae bacteriophage MM1] emb CAB96622.1 (AJ400631) hypothetical protein [Streptococcus pneumoniae] Length = 303
SeqID 562	SA-1442.2	Contig139 (81115-82089 p)	86	Identities = 254/321 (79%), Positives = 286/321 (88%), Gaps = 1/321 (0%) emb CAB96620.1 (AJ400630) hypothetical protein [Streptococcus pneumoniae bacteriophage MM1] emb CAB96623.1 (AJ400631) hypothetical protein [Streptococcus pneumoniae] Length = 325
SeqID 563	SA-1445.2	Contig139 (80228-81118 p)	68	Identities = 165/291 (56%), Positives = 208/291 (70%), Gaps = 2/291 (0%) sp O06973 YVCJ_BACSU HYPOTHETICAL 33.9 KD PROTEIN IN CRH-TRXB INTERGENIC REGION pir H70031 conserved hypothetical protein yvcJ - Bacillus subtilis emb CAB08057.1 (Z94043) hypothetical protein [Bacillus subtilis] emb CAB15482.1 (Z99121) similar to hypothetical proteins [Bacillus subtilis] Length = 295

SeqID 564	SA-1446.1	Contig139 (79661-80074 m)	36	Identities = 36/134 (26%), Positives = 57/134 (41%), Gaps = 7/134 (5%) emb CAB62728.1 (AL133423) hypothetical protein SC4A7.24c [Streptomyces coelicolor A3(2)] Length = 139
SeqID 565	SA-1447.1	Contig139 (78432-79415 m)	46	Identities = 97/305 (31%), Positives = 160/305 (51%), Gaps = 10/305 (3%) sp P33022 YEIK_ECOLI HYPOTHETICAL 33.7 KD PROTEIN IN NFO-FRUA INTERGENIC REGION pir A64985 hypothetical 33.7 kD protein in nfo-fruA intergenic region - Escherichia coli (strain K-12) gb AAA60514.1 (U00007) yeik [Escherichia coli] gb AAC75223.1 (AE000305) orf, hypothetical protein [Escherichia coli K12] prf 2014253BH yeik gene [Escherichia coli] Length = 313
SeqID 566	SA-1448.1	Contig139 (77878-78435 m)	38	Identities = 39/135 (28%), Positives = 72/135 (52%), Gaps = 4/135 (2%) ref NP_053012.1 hypothetical protein [Plasmid pNZ4000] gb AAD40355.1 (AF036485) hypothetical protein [Plasmid pNZ4000] Length = 200
SeqID 567	SA-1449.1	Contig139 (76492-77838 p)	69	Identities = 252/442 (57%), Positives = 316/442 (71%), Gaps = 15/442 (3%) dbj BAB05415.1 (AP001512) asparaginyl-tRNA synthetase [Bacillus halodurans] Length = 430
SeqID 568	SA-145.1	Contig138 (87488-87769 p)	No Hits found	
SeqID 569	SA-1450.1	Contig139 (75278-76471 p)	75	Identities = 270/391 (69%), Positives = 314/391 (80%) gb AAF12702.1 AF035157_1 (AF035157) aspartate aminotransferase [Lactococcus lactis] Length = 393
SeqID 570	SA-1451.1	Contig139 (72727-75192 p)	47	Identities = 251/927 (27%), Positives = 398/927 (42%), Gaps = 145/927 (15%) dbj BAB05410.1 (AP001512) ATP-dependent DNA helicase [Bacillus halodurans] Length = 942
SeqID 571	SA-1453.2	Contig139 (71813-72577 p)	69	Identities = 162/254 (63%), Positives = 188/254 (73%), Gaps = 2/254 (0%) gb AAC48769.1 (U71200) acetoin reductase [Bos taurus] Length = 257

SeqID 572	SA-1456.1	Contig95 (12824-13120 m)	60	Identities = 53/112 (47%), Positives = 78/112 (69%), Gaps = 1/112 (0%) gb AAC23746.1 (AF052209) competence protein [Streptococcus pneumoniae] Length = 266
SeqID 573	SA-1457.1	Contig95 (10627-12858 p)	29	Identities = 142/468 (30%), Positives = 220/468 (46%), Gaps = 62/468 (13%) pir F82995 glutamate--cysteine ligase PA5203 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG08588.1 AE004933_4 (AE004933) glutamate--cysteine ligase [Pseudomonas aeruginosa] Length = 527
SeqID 574	SA-1458.1	Contig95 (9882-10556 p)	64	Identities = 107/223 (47%), Positives = 148/223 (65%), Gaps = 7/223 (3%) pir H81283 helix-turn-helix containing protein Cj1387c [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73814.1 (AL139078) helix-turn-helix containing protein [Campylobacter jejuni] Length = 218
SeqID 575	SA-146.1	Contig138 (86863-87177 m)	No Hits found	
SeqID 576	SA-1460.1	Contig95 (8498-9520 m)	77	Identities = 225/341 (65%), Positives = 279/341 (80%) pir A43577 regulatory protein pfoR - Clostridium perfringens Length = 343
SeqID 577	SA-1462.1	Contig95 (6835-8127 m)	87	Identities = 320/427 (74%), Positives = 378/427 (87%) pir A42280 adenylosuccinate synthase (EC 6.3.4.4) purA - Bacillus subtilis dbj BAA05174.1 (D26185) adenylosuccinate synthetase [Bacillus subtilis] emb CAB16079.1 (Z99124) adenylosuccinate synthetase [Bacillus subtilis] Length = 430
SeqID 578	SA-1463.1	Contig95 (6339-6812 p)	No Hits found	
SeqID 579	SA-1465.1	Contig95 (5563-6180 m)	No Hits found	
SeqID 580	SA-1466.1	Contig95 (3956-5395 m)	55	Identities = 184/465 (39%), Positives = 279/465 (59%), Gaps = 25/465 (5%) sp P39301 SGAT_ECOLI_PUTATIVE TRANSPORT PROTEIN SGAT pir D65230 hypothetical 52.9 kD protein in aidB-rpsF intergenic region - Escherichia coli (strain K-12) gb AAC77150.1 (AE000491) orf, hypothetical protein [Escherichia coli K12] Length = 484

SeqID 581	SA-1467.1	Contig95 (3650-3928 m)	64	Identities = 42/95 (44%), Positives = 65/95 (68%), Gaps = 1/95 (1%) gb AAG34743.1 AE000033_7 (AE000033) similar to PTS system: EII ^B [Mycoplasma pneumoniae] Length = 95
SeqID 582	SA-1468.1	Contig95 (3098-3583 m)	58	Identities = 64/150 (42%), Positives = 97/150 (64%), Gaps = 2/150 (1%) sp P39303 PTXA_ECOLI_UNKNOWN_PENTITOL_PHOSPHOTRANSFERASE_ENZYME_II, A COMPONENT pir F65230 hypothetical phosphotransferase enzyme II - Escherichia coli (strain K-12) gb AAC77152.1 (AE000491) putative PTS system enzyme II A component [Escherichia coli K12] Length = 154
SeqID 583	SA-1469.1	Contig95 (2320-2985 m)	60	Identities = 108/217 (49%), Positives = 141/217 (64%), Gaps = 3/217 (1%) sp P39304 SGAH_ECOLI_PROBABLE_HEXULOSE-6-PHOSPHATE SYNTHASE (HUMPS) (D-ARABINO 3-HEXULOSE 6-PHOSPHATE FORMALDEHYDE LYASE) pir S56421 hypothetical 23.6K protein (aidB-rpsF intergenic region) - Escherichia coli gb AAA97092.1 (U14003) ORF_o216 [Escherichia coli] gb AAC77153.1 (AE000491) probable hexulose-6-phosphate synthase [Escherichia coli K12] Length = 216
SeqID 584	SA-147.1	Contig138 (85926-86402 m)	37	Identities = 42/153 (27%), Positives = 71/153 (45%), Gaps = 7/153 (4%) emb CAB76310.1 (AL158060) putative acetyltransferase. [Streptomyces coelicolor A3(2)] Length = 183
SeqID 585	SA-1470.1	Contig95 (1453-2316 m)	68	Identities = 147/285 (51%), Positives = 198/285 (68%), Gaps = 9/285 (3%) sp P39305 SGAU_ECOLI_PUTATIVE_HEXULOSE-6-PHOSPHATE ISOMERASE (HUMPI) pir S56422 hypothetical 32K protein (aidB-rpsF intergenic region) - Escherichia coli gb AAA97093.1 (U14003) ORF_o284 [Escherichia coli] gb AAC77154.1 (AE000491) putative hexulose-6-phosphate isomerase [Escherichia coli K12] Length = 284

SeqID 586	SA-1471.1	Contig95 (735-1451 m)	73	Identities = 143/229 (62%), Positives = 176/229 (76%), Gaps = 2/229 (0%) gb AAD45716.1 AF160811_4 (AF160811) L-ribulose 5-phosphate 4-epimerase [Bacillus stearothermophilus] Length = 228
SeqID 587	SA-1472.1	Contig95 (77-724 m)	60	Identities = 124/214 (57%), Positives = 157/214 (72%) sp Q58370 TAL_METJA TRANSALDOLASE-LIKE PROTEIN pir H64419 transaldolase (EC 2.2.1.2) - Methanococcus jannaschii gb AAB98962.1 (U67539) transaldolase [Methanococcus jannaschii] Length = 217
SeqID 588	SA-1473.2	Contig94 (12280-12639 p)	No Hits found	
SeqID 589	SA-1474.1	Contig94 (11451-11573 m)	No Hits found	
SeqID 590	SA-1475.1	Contig94 (10797-11117 m)	No Hits found	
SeqID 591	SA-1477.1	Contig94 (9456-9902 m)	48	Identities = 59/185 (31%), Positives = 96/185 (51%), Gaps = 12/185 (6%) pir A82466 hypothetical protein VCA0380 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF96286.1 (AE004374) hypothetical protein [Vibrio cholerae] Length = 201
SeqID 592	SA-1478.1	Contig94 (8817-9236 m)	62	Identities = 49/147 (33%), Positives = 98/147 (66%), Gaps = 2/147 (1%) pir A69849 hypothetical protein yjdF - Bacillus subtilis emb CAB13060.1 (Z99110) yjdF [Bacillus subtilis] Length = 160
SeqID 593	SA-1479.1	Contig94 (8138-8617 m)	58	Identities = 61/152 (40%), Positives = 95/152 (62%) gb AAD50427.1 AF161700_2 (AF161700) ComX1 [Streptococcus pneumoniae] gb AAD50429.1 AF161701_2 (AF161701) ComX2 [Streptococcus pneumoniae] Length = 159
SeqID 594	SA-148.1	Contig138 (85248-85790 m)	46	Identities = 55/175 (31%), Positives = 93/175 (52%), Gaps = 6/175 (3%) sp P05332 YP20_BACLI HYPOTHETICAL P20 PROTEIN pir S00875 hypothetical protein P20 - Bacillus licheniformis emb CAA30415.1 (X07542) P20 (AA 1-178) [Bacillus licheniformis] Length = 178

SeqID 595	SA-1480.1	Contig94 (7324-8016 m)	41	Identities = 57/226 (25%), Positives = 97/226 (42%), Gaps = 22/226 (9%) ref NP_014926.1 Yor283wp [Saccharomyces cerevisiae] pir S67185 hypothetical protein YOR283w - yeast (Saccharomyces cerevisiae) emb CAA61787.1 (X89633) hypothetical phosphoglycerate mutase [Saccharomyces cerevisiae] emb CAA99510.1 (Z75191) ORF YOR283w [Saccharomyces cerevisiae] Length = 230
SeqID 596	SA-1481.1	Contig94 (6575-7249 m)	50	Identities = 108/169 (63%), Positives = 139/169 (81%) gb AAD00280.1 (U78599) putative D,D-carboxypeptidase [Streptococcus mutans] Length = 173
SeqID 597	SA-1482.1	Contig94 (6003-6425 m)	48	Identities = 65/153 (42%), Positives = 94/153 (60%), Gaps = 4/153 (2%) gb AAD00279.1 (U78599) putative N-acetyl-muramidase [Streptococcus mutans] Length = 158
SeqID 598	SA-1483.1	Contig94 (4826-5860 m)	80	Identities = 233/344 (67%), Positives = 294/344 (84%) sp O06940 HRCA_STRMU HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR HRCA Length = 344
SeqID 599	SA-1484.1	Contig94 (4251-4784 m)	78	Identities = 129/188 (68%), Positives = 153/188 (80%), Gaps = 2/188 (1%) pir S39341 grpE protein - Lactococcus lactis Length = 190
SeqID 600	SA-1486.1	Contig94 (2241-4070 m)	97	Identities = 609/609 (100%), Positives = 609/609 (100%) sp P95693 DNAK_STRAG DNAK PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70) gb AAB39219.1 (U72719) heat shock protein 70 [Streptococcus agalactiae] Length = 609
SeqID 601	SA-1487.1	Contig94 (813-1952 m)	93	Identities = 328/377 (87%), Positives = 356/377 (94%), Gaps = 1/377 (0%) dbj BAB16032.1 (AB030809) Streptococcus pneumoniae DnaJ protein homologue [Streptococcus pyogenes] Length = 378
SeqID 602	SA-1488.1	Contig94 (495-827 p)	No Hits found	
SeqID 603	SA-1489.1	Contig94 (1-522 p)	79	Identities = 122/164 (74%), Positives = 138/164 (83%) sp Q07211 SCRK_STRMU FRUCTOKINASE dbj BAA02467.1 (D13175) fructokinase [Streptococcus mutans] Length = 293

SeqID 604	SA-149.1	Contig138 (84525-85205 m)	64	Identities = 110/230 (47%), Positives = 154/230 (66%), Gaps = 3/230 (1%) ref NP_012308.1 Yir042cp [Saccharomyces cerevisiae] sp P40586 YIW2_YEAST_HYPOTHETICAL 27.4 KD PROTEIN IN HYR1 3 REGION pir S50347 hypothetical protein YIR042c - yeast (Saccharomyces cerevisiae) emb CAA87001.1 (Z46902) unknown [Saccharomyces cerevisiae] Length = 236
SeqID 605	SA-1490.2	Contig134 (49677-50102 p)	82	Identities = 118/139 (84%), Positives = 129/139 (91%) sp P36254 RL11_STACA 50S RIBOSOMAL PROTEIN L11 pir S38871 ribosomal protein L11 - Staphylococcus carnosus emb CAA53739.1 (X76134) L11 protein [Staphylococcus carnosus] Length = 140
SeqID 606	SA-1491.1	Contig134 (48084-49466 p)	43	Identities = 118/405 (29%), Positives = 199/405 (49%), Gaps = 9/405 (2%) pir F70705 hypothetical protein Rv2333c - Mycobacterium tuberculosis (strain H37RV) emb CAB02058.1 (Z79702) hypothetical protein Rv2333c [Mycobacterium tuberculosis] Length = 537
SeqID 607	SA-1492.1	Contig134 (46862-48076 p)	66	Identities = 201/395 (50%), Positives = 274/395 (68%), Gaps = 5/395 (1%) gb AAF36227.1 AF168363_2 (AF168363) aminoacylase [Lactococcus lactis] Length = 398
SeqID 608	SA-1493.1	Contig134 (45698-46582 m)	40	Identities = 72/259 (27%), Positives = 127/259 (48%), Gaps = 7/259 (2%) ref NP_043632.1 former trsE (rbcR homolog) [Odontella sinensis] sp P49518 YC30_ODOSI_HYPOTHETICAL TRANSCRIPTIONAL REGULATOR YCF30 pir S78291 probable transcription regulator ycf30 - Odontella sinensis chloroplast emb CAA91664.1 (Z67753) former trsE (rbcR homolog) [Odontella sinensis] Length = 309
SeqID 609	SA-1494.1	Contig134 (45284-45646 m)	No Hits found	

SeqID 610	SA-1495.1	Contig134 (43540-45255 p)	56	Identities = 203/580 (35%), Positives = 325/580 (56%), Gaps = 50/580 (8%) pir E64556 para-aminobenzoate synthetase - Helicobacter pylori (strain 26695) gb AAD07357.1 (AE000547) para-aminobenzoate synthetase (pabB) [Helicobacter pylori 26695] Length = 559
SeqID 611	SA-1498.1	Contig134 (41017-43458 p)	62	Identities = 353/828 (42%), Positives = 511/828 (61%), Gaps = 70/828 (8%) sp P21458 SP3E_BACSU STAGE III SPORULATION PROTEIN E pir S09411 DNA translocase sp IIE - Bacillus subtilis emb CAB13553.1 (Z99112) DNA translocase [Bacillus subtilis] Length = 787
SeqID 612	SA-1499.1	Contig134 (40039-40842 m)	44	Identities = 88/224 (39%), Positives = 123/224 (54%), Gaps = 46/224 (20%) pir T41399 probable cyclophilin-related peptidyl prolyl cis-trans isomerase - fission yeast (Schizosaccharomyces pombe) emb CAA19257.1 (AL023704) putative Cyclophilin-type peptidyl-prolyl cis-trans isomerase protein [Schizosaccharomyces pombe] Length = 610
SeqID 613	SA-15.1	Contig137 (29069-31870 p)	22	Identities = 152/405 (37%), Positives = 214/405 (52%), Gaps = 26/405 (6%) pir A43607 cell surface antigen SpaA precursor - Streptococcus sobrinus (strain MT3791) dbj BAA14368.1 (D90354) surface protein antigen precursor [Streptococcus sobrinus] Length = 1566
SeqID 614	SA-150.1	Contig138 (83318-84265 p)	No Hits found	
SeqID 615	SA-1500.1	Contig134 (39154-39987 p)	90	Identities = 224/275 (81%), Positives = 255/275 (92%) gb AAD56938.1 AF180520_3 (AF180520) integral membrane protein MtsC [Streptococcus pyogenes] Length = 284
SeqID 616	SA-1501.1	Contig134 (38436-39152 p)	72	Identities = 143/238 (60%), Positives = 186/238 (78%), Gaps = 2/238 (0%) gb AAD56937.1 AF180520_2 (AF180520) ATP-binding protein MtsB [Streptococcus pyogenes] Length = 241

SeqID 617	SA-1502.2	Contig134 (37339-38265 p)	86	Identities = 240/308 (77%), Positives = 277/308 (89%), Gaps = 1/308 (0%) gb AAD56936.1 AF180520_1 (AF180520) lipoprotein MtsA [Streptococcus pyogenes] Length = 310
SeqID 618	SA-1503.2	Contig108 (5713-10425 m)	66	Identities = 795/1596 (49%), Positives = 1056/1596 (65%), Gaps = 39/1596 (2%) gb AAG09771.1 AF243528_1 (AF243528) cell envelope proteinase [Streptococcus thermophilus] Length = 1585
SeqID 619	SA-1504.1	Contig108 (4794-5480 m)	50	Identities = 87/220 (39%), Positives = 124/220 (55%), Gaps = 4/220 (1%) dbj BAB05663.1 (AP001513) two-component response regulator [Bacillus halodurans] Length = 229
SeqID 620	SA-1506.1	Contig108 (3409-4665 m)	40	Identities = 108/379 (28%), Positives = 193/379 (50%), Gaps = 33/379 (8%) pir D70045 two-component sensor histidine kinase homolog yvqB - Bacillus subtilis emb CAB15292.1 (Z99120) similar to two-component sensor histidine kinase [YvqA] [Bacillus subtilis] emb CAA11751.1 (AJ223978) putative sensory protein kinase, YvqB [Bacillus subtilis] Length = 451
SeqID 621	SA-1507.1	Contig108 (2708-3334 m)	33	Identities = 38/144 (26%), Positives = 69/144 (47%), Gaps = 10/144 (6%) ref XP_005848.1 KIAA1074 protein [Homo sapiens] Length = 1709
SeqID 622	SA-151.1	Contig138 (82830-83321 p)	42	Identities = 40/141 (28%), Positives = 75/141 (52%), Gaps = 3/141 (2%) dbj BAB05334.1 (AP001512) RNA polymerase sigma factor Y [Bacillus halodurans] Length = 176
SeqID 623	SA-1511.1	Contig108 (120-2621 p)	78	Identities = 569/835 (68%), Positives = 666/835 (79%), Gaps = 42/835 (5%) sp P36430 SYL_BACSU LEUCYL-TRNA SYNTHETASE (LEUCINE--TRNA LIGASE) (LEURS) pir D69650 leucine--trna ligase (EC 6.1.1.4) - Bacillus subtilis gb AAC00259.1 (AF008220) leucine tRNA synthetase [Bacillus subtilis] emb CAB15010.1 (Z99119) leucyl-tRNA synthetase [Bacillus subtilis] Length = 804

SeqID 624	SA-1512.2	Contig138 (776-1330 m)	71	Identities = 107/185 (57%), Positives = 140/185 (74%), Gaps = 3/185 (1%) emb CAB91549.1 (AJ249134) hypothetical protein [Lactococcus lactis] Length = 185
SeqID 625	SA-1513.1	Contig138 (1407-2072 m)	51	Identities = 76/230 (33%), Positives = 118/230 (51%), Gaps = 11/230 (4%) sp P39147 CMF3_BACSU COMF OPERON PROTEIN 3 pir J40389 hypothetical protein F3 - Bacillus subtilis pir S77621 late competence gene comFC - Bacillus subtilis emb CAA79228.1 (Z18629) F3 [Bacillus subtilis] gb AAC44942.1 (U56901) involved in transformation [Bacillus subtilis] emb CAB15562.1 (Z99122) late competence gene [Bacillus subtilis] Length = 229
SeqID 626	SA-1514.1	Contig138 (2838-3017 p)	No Hits found	
SeqID 627	SA-1515.1	Contig138 (2072-3361 m)	52	Identities = 160/405 (39%), Positives = 228/405 (55%), Gaps = 20/405 (4%) sp P39145 CMF1_BACSU COMF OPERON PROTEIN 1 pir J40387 hypothetical protein F1 - Bacillus subtilis pir G69602 late competence protein required for DNA uptake comFA - Bacillus subtilis emb CAA79226.1 (Z18629) F1 [Bacillus subtilis] gb AAC44940.1 (U56901) involved in transformation [Bacillus subtilis] emb CAB15564.1 (Z99122) late competence protein [Bacillus subtilis] Length = 463
SeqID 628	SA-1516.1	Contig138 (3417-4061 p)	58	Identities = 96/204 (47%), Positives = 127/204 (62%) dbj BAB07349.1 (AP001519) unknown conserved protein [Bacillus halodurans] Length = 213
SeqID 629	SA-1517.1	Contig138 (4152-5078 p)	88	Identities = 239/306 (78%), Positives = 274/306 (89%) dbj BAA88310.1 (AB028865) O-acetylserine lyase [Streptococcus suis] Length = 308
SeqID 630	SA-1518.1	Contig138 (5230-6324 m)	73	Identities = 199/363 (54%), Positives = 272/363 (74%), Gaps = 2/363 (0%) sp P32816 GLDA_BACST GLYCEROL DEHYDROGENASE (GLDH) pir JQ1474 glycerol dehydrogenase (EC 1.1.1.6) [validated] - Bacillus stearothermophilus gb AAA22477.1 (M65289) glycerol dehydrogenase [Bacillus stearothermophilus] Length = 370

SeqID 631	SA-1519.1	Contig138 (6386-7054 m)	57	Identities = 90/214 (42%), Positives = 131/214 (61%), Gaps = 1/214 (0%) emb CAA05516.1 (AJ002527) OrfX [Clostridium beijerinckii] Length = 226
SeqID 632	SA-152.1	Contig138 (81640-82248 p)	42	Identities = 48/182 (26%), Positives = 91/182 (49%), Gaps = 12/182 (6%) pir B70391 transcription regulator TetR/AcrR family - Aquifex aeolicus gb AAC07123.1 (AE000721) transcriptional regulator (TetR/AcrR family) [Aquifex aeolicus] Length = 179
SeqID 633	SA-1520.3	Contig138 (7064-9448 m)	67	Identities = 414/805 (51%), Positives = 555/805 (68%), Gaps = 14/805 (1%) sp P75793 PFLF_ECOLI PUTATIVE FORMATE ACETYLTRANSFERASE 3 (PYRUVATE FORMATE-LYASE 3) pir G64819 probable formate C-acetyltransferase (EC 2.3.1.54) - Escherichia coli gb AAC73910.1 (AE000184) putative formate acetyltransferase [Escherichia coli K12] dbj BAA35511.1 (D90720) Formate acetyltransferase 2 (EC 2.3.1.54) (pyruvate formate-lyase 2). [Escherichia coli] Length = 810
SeqID 634	SA-1521.1	Contig135 (23461-23691 p)	No Hits found	
SeqID 635	SA-1522.1	Contig135 (22356-23378 p)	84	Identities = 259/340 (76%), Positives = 294/340 (86%) gb AAF34406.1 AF164204_1 (AF164204) branched-chain amino acid aminotransferase [Lactococcus lactis subsp. cremoris] gb AAF64593.1 AF169649_1 (AF169649) branched-chain aminotransferase lve [Lactococcus lactis] Length = 340
SeqID 636	SA-1523.1	Contig135 (19784-22243 p)	87	Identities = 640/821 (77%), Positives = 724/821 (87%), Gaps = 5/821 (0%) gb AAD34369.1 AF129764_3 (AF129764) ParC [Streptococcus mitis] Length = 823
SeqID 637	SA-1524.1	Contig135 (17701-19650 p)	94	Identities = 574/649 (88%), Positives = 617/649 (94%), Gaps = 2/649 (0%) emb CAA91550.2 (Z67739) DNA topoisomerase IV [Streptococcus pneumoniae] Length = 647

SeqID 638	SA-1525.1	Contig135 (16949-17587 m)	67	Identities = 134/213 (62%), Positives = 169/213 (78%) sp Q54916 YPAE_STRPN_HYPOTHETICAL_23.0_KD_PROTEIN IN_PARE_5_REGION_(ORF2)_emb CAA91549.1 (Z67739) unidentified [Streptococcus pneumoniae] Length = 213
SeqID 639	SA-1526.1	Contig135 (16230-16883 p)	99	Identities = 217/217 (100%), Positives = 217/217 (100%) sp Q9XDS8 UNG_STRAG_URACIL-DNA_GLYCOSYLASE_(UDG) dbj BAA82292.1 (AB028896) Uracil DNA glycosylase [Streptococcus agalactiae] Length = 217
SeqID 640	SA-1527.1	Contig135 (15646-16131 p)	96	Identities = 161/161 (100%), Positives = 161/161 (100%) dbj BAA82291.1 (AB028896) ORF1 [Streptococcus agalactiae] Length = 161
SeqID 641	SA-1528.1	Contig135 (14291-15532 p)	98	Identities = 413/413 (100%), Positives = 413/413 (100%) dbj BAA82290.1 (AB028896) NeuA [Streptococcus agalactiae] Length = 413
SeqID 642	SA-153.1	Contig138 (80519-81616 p)	30	Identities = 64/187 (34%), Positives = 112/187 (59%), Gaps = 1/187 (0%) pir T36287 probable integral membrane protein - Streptomyces coelicolor emb CAB42664.1 (AL049819) putative integral membrane protein [Streptomyces coelicolor A3(2)] Length = 266
SeqID 643	SA-1530.1	Contig135 (13651-14280 p)	98	Identities = 209/209 (100%), Positives = 209/209 (100%) dbj BAA82289.1 (AB028896) NeuD [Streptococcus agalactiae] Length = 209
SeqID 644	SA-1531.1	Contig135 (12500-13654 p)	96	Identities = 384/384 (100%), Positives = 384/384 (100%) dbj BAA82288.1 (AB028896) NeuC [Streptococcus agalactiae] Length = 384
SeqID 645	SA-1532.1	Contig135 (11398-12423 p)	97	Identities = 341/341 (100%), Positives = 341/341 (100%) dbj BAA33753.1 (AB017355) neuB [Streptococcus agalactiae] dbj BAA82287.1 (AB028896) NeuB [Streptococcus agalactiae] Length = 341

SeqID 646	SA-1533.1	Contig135 (9998-11398 p)	99	<p>Identities = 466/466 (100%), Positives = 466/466 (100%) pir T44650 capsular polysaccharide repeat unit transporter cpsM [imported] - Streptococcus agalactiae dbj BAA33752.1 (AB017355) capsular polysaccharide [Streptococcus agalactiae] dbj BAA82286.1 (AB028896) CpsIA [Streptococcus agalactiae] gb AAD53073.1 AF163833_13 (AF163833) CpsM [Streptococcus agalactiae] Length = 466</p>
SeqID 647	SA-1535.2	Contig106 (11312-12010 m)	64	<p>Identities = 87/228 (38%), Positives = 150/228 (65%), Gaps = 1/228 (0%) sp P54591 YHCG_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 1 IN GLPD-CSPB INTERGENIC REGION pir C69822 glycine betaine/L-proline transport homolog yhcG - Bacillus subtilis emb CAA65690.1 (X96983) hypothetical protein [Bacillus subtilis] emb CAB12735.1 (Z99108) similar to glycine betaine/L-proline transport [Bacillus subtilis] Length = 232</p>
SeqID 648	SA-1536.1	Contig106 (10518-11300 m)	37	<p>Identities = 49/208 (23%), Positives = 102/208 (48%), Gaps = 20/208 (9%) gb AAA29909.1 (M74170) ORF 3 [Schistosoma mansoni] Length = 393</p>
SeqID 649	SA-1537.1	Contig106 (9925-10482 p)	40	<p>Identities = 43/174 (24%), Positives = 84/174 (47%), Gaps = 9/174 (5%) pir F64934 hypothetical protein b1750 - Escherichia coli (strain K-12) gb AAC74820.1 (AE000270) orf, hypothetical protein [Escherichia coli K12] Length = 252</p>
SeqID 650	SA-1538.1	Contig106 (8523-9152 m)	39	<p>Identities = 63/135 (46%), Positives = 83/135 (60%), Gaps = 6/135 (4%) sp P26840 MATA_BACSH PROBABLE MACROLIDE ACETYLTRANSFERASE Length = 180</p>
SeqID 651	SA-1539.1	Contig106 (7434-8402 p)	No Hits found	
SeqID 652	SA-1542.1	Contig106 (5238-5555 m)	49	<p>Identities = 37/92 (40%), Positives = 52/92 (56%), Gaps = 10/92 (10%) emb CAA59764.1 (X85757) unknown [Saccharomyces cerevisiae] Length = 133</p>

SeqID 653	SA-1543.1	Contig106 (4507-7299 p)	65	Identities = 410/922 (44%), Positives = 609/922 (65%), Gaps = 22/922 (2%) pir C69069 cation-transporting P-ATPase PacL - Methanobacterium thermoautotrophicum (strain Delta H) gb AAB85991.1 (AE000912) cation-transporting P-ATPase PacL [Methanobacterium thermoautotrophicum] Length = 910
SeqID 654	SA-1544.1	Contig106 (3921-4355 p)	67	Identities = 85/132 (64%), Positives = 105/132 (79%) emb CAB46979.1 (AJ243482) CSRA protein [Enterococcus faecalis] Length = 168
SeqID 655	SA-1545.1	Contig106 (3451-3921 p)	46	Identities = 48/152 (31%), Positives = 75/152 (48%), Gaps = 1/152 (0%) dbj BAB05127.1 (AP001511) unknown [Bacillus halodurans] Length = 156
SeqID 656	SA-1546.1	Contig106 (2795-3448 p)	66	Identities = 108/211 (51%), Positives = 149/211 (70%) dbj BAB06554.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 215
SeqID 657	SA-1547.1	Contig106 (26-2689 p)	14	Identities = 76/279 (27%), Positives = 133/279 (47%), Gaps = 4/279 (1%) spi P25146 INLA_LISMO INTERNALIN A PRECURSOR pir S37387 internalin A precursor - Listeria monocytogenes Length = 800
SeqID 658	SA-155.1	Contig138 (79800-80516 p)	54	Identities = 86/232 (37%), Positives = 137/232 (58%), Gaps = 3/232 (1%) gb AAF50837.1 (AE003568) CG1718 gene product [Drosophila melanogaster] Length = 1713
SeqID 659	SA-1551.2	Contig139 (128987-131698 p)	6	Identities = 41/110 (37%), Positives = 56/110 (50%), Gaps = 4/110 (3%) gb AAD33086.1 AF071083_1 (AF071083) fibronectin- binding protein I [Streptococcus pyogenes] Length = 1161
SeqID 660	SA-1552.1	Contig139 (131737-131979 p)	No Hits found	
SeqID 661	SA-1554.1	Contig139 (132099-132542 p)	42	Identities = 58/135 (42%), Positives = 86/135 (62%), Gaps = 2/135 (1%) gb AAC13546.1 (AF019629) putative fimbria- associated protein [Actinomyces naeslundii] Length = 365

SeqID 662	SA-1555.1	Contig139 (132758-133396 p)	19	Identities = 30/96 (31%), Positives = 46/96 (47%), Gaps = 9/96 (9%) gb AAD33086.1 AF071083_1 (AF071083) fibronectin-binding protein I [Streptococcus pyogenes] Length = 1161
SeqID 663	SA-1556.1	Contig139 (133726-133872 p)	No Hits found	
SeqID 664	SA-1558.1	Contig139 (134494-134670 p)	46	Identities = 23/67 (34%), Positives = 37/67 (54%), Gaps = 2/67 (2%) gb AAB00100.1 (L36660) unknown [Streptococcus pneumoniae] emb CAA84075.1 (Z34303) hypothetical protein [Streptococcus pneumoniae] Length = 194
SeqID 665	SA-156.1	Contig138 (78853-79536 p)	38	Identities = 45/205 (21%), Positives = 93/205 (44%), Gaps = 10/205 (4%) dbj BAB04126.1 (AP001508) unknown conserved protein in others [Bacillus halodurans] Length = 228
SeqID 666	SA-1560.1	Contig139 (135988-136284 p)	No Hits found	
SeqID 667	SA-1562.1	Contig139 (136414-137496 p)	31	Identities = 72/279 (25%), Positives = 125/279 (43%), Gaps = 25/279 (8%) dbj BAA24012.1 (AB009635) Fmt [Staphylococcus aureus] Length = 397
SeqID 668	SA-1563.1	Contig139 (137702-138688 p)	58	Identities = 141/331 (42%), Positives = 201/331 (60%), Gaps = 34/331 (10%) pir G75548 ABC transporter, ATP-binding protein - Deinococcus radiodurans (strain R1) gb AAF09790.1 AE001882_8 (AE001882) ABC transporter, ATP-binding protein [Deinococcus radiodurans] Length = 354
SeqID 669	SA-1564.1	Contig139 (138691-139509 p)	42	Identities = 60/271 (22%), Positives = 118/271 (43%), Gaps = 13/271 (4%) emb CAB88837.1 (AL353832) putative integral membrane protein. [Streptomyces coelicolor A3(2)] Length = 268

SeqID 670	SA-1565.2	Contig139 (139511-140296 p)	44	Identities = 68/264 (25%), Positives = 123/264 (45%), Gaps = 10/264 (3%) emb CAB88836.1 (AL353832) putative integral membrane transport protein. [Streptomyces coelicolor A3(2)] Length = 295
SeqID 671	SA-1567.3	Contig128 (12896-13207 m)	41	Identities = 25/84 (29%), Positives = 45/84 (52%), Gaps = 6/84 (7%) gb AAD31042.1 AF144646_1 (AF144646) heat shock protein 70 [Crassostrea gigas] Length = 659
SeqID 672	SA-1568.3	Contig128 (13333-14211 p)	67	Identities = 175/280 (62%), Positives = 218/280 (77%), Gaps = 9/280 (3%) sp O07874 RNH3_STRPN RIBONUCLEASE HIII (RNASE HIII) gb AAC45437.1 (U93576) ribonuclease HII [Streptococcus pneumoniae] Length = 290
SeqID 673	SA-157.1	Contig138 (78521-78832 p)	72	Identities = 67/133 (50%), Positives = 86/133 (64%), Gaps = 16/133 (12%) dbj BAB04125.1 (AP001508) unknown conserved protein in others [Bacillus halodurans] Length = 174
SeqID 674	SA-1570.1	Contig128 (14227-14820 p)	71	Identities = 120/203 (59%), Positives = 145/203 (71%), Gaps = 9/203 (4%) sp O07344 LEP_STRPN SIGNAL PEPTIDASE I (SPASE I) (LEADER PEPTIDASE I) gb AAB69116.1 (U90721) signal peptidase I [Streptococcus pneumoniae] Length = 204
SeqID 675	SA-1571.2	Contig128 (14949-17369 p)	73	Identities = 454/835 (54%), Positives = 600/835 (71%), Gaps = 37/835 (4%) gb AAK05838.1 AE006404_5 (AE006404) exodeoxyribonuclease V alpha chain (EC 3.1.11.5) [Lactococcus lactis subsp. lactis] Length = 834
SeqID 676	SA-1573.3	Contig128 (17483-17965 p)	42	Identities = 47/173 (27%), Positives = 76/173 (43%), Gaps = 17/173 (9%) emb CAA72923.1 (Y12234) hypothetical protein [Enterococcus faecalis] Length = 169

SeqID 677	SA-1574.1	Contig128 (18036-19055 m)	55	Identities = 136/349 (38%), Positives = 210/349 (59%), Gaps = 14/349 (4%) pir G82093 DNA-damage-inducible protein P VC2287 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAAF95431.1 (AE004300) DNA-damage-inducible protein P [Vibrio cholerae] Length = 360
SeqID 678	SA-1577.1	Contig128 (19315-21627 p)	95	Identities = 708/770 (91%), Positives = 750/770 (96%) spiQ59334 PFL_STRMU FORMATE ACETYLTRANSFERASE (PYRUVATE FORMATE-LYASE) dbj BAA09085.1 (D50491) Pyruvate formate-lyase [Streptococcus mutans] Length = 775
SeqID 679	SA-1578.1	Contig128 (21728-22102 p)	55	Identities = 53/124 (42%), Positives = 76/124 (60%), Gaps = 2/124 (1%) sp Q46604 FMNB_DESVM FMN-BINDING PROTEIN pdb 1FLM A Chain A, Dimer Of Fmn-Binding Protein From Desulfovibrio Vulgaris (Miyazaki F) pdb 1FLM B Chain B, Dimer Of Fmn-Binding Protein From Desulfovibrio Vulgaris (Miyazaki F) pdb 1AXJ Fmn-Binding Protein From Desulfovibrio Vulgaris (Miyazaki F), Nmr, 20 Structures dbj BAA25177.1 (D21804) FMN-binding protein [Desulfovibrio vulgaris] Length = 122
SeqID 680	SA-1579.1	Contig128 (22137-23081 m)	44	Identities = 88/323 (27%), Positives = 142/323 (43%), Gaps = 51/323 (15%) emb CAC16441.1 (AL450165) putative esterase [Streptomyces coelicolor] Length = 353
SeqID 681	SA-1580.1	Contig128 (23063-23818 m)	55	Identities = 92/240 (38%), Positives = 143/240 (59%), Gaps = 11/240 (4%) gb AAD37110.1 (AF112358) C3-degrading proteinase [Streptococcus pneumoniae] Length = 241
SeqID 682	SA-1582.1	Contig128 (23943-24839 p)	21	Identities = 43/180 (23%), Positives = 71/180 (38%), Gaps = 16/180 (8%) dbj BAB07173.1 (AP001518) unknown conserved protein [Bacillus halodurans] Length = 273
SeqID 683	SA-1584.3	Contig128 (24918-25745 m)	77	Identities = 186/290 (64%), Positives = 228/290 (78%), Gaps = 10/290 (3%) gb AAK06146.1 AE006435_4 (AE006435) transporter [Lactococcus lactis subsp. lactis] Length = 289

SeqID 684	SA-1585.3	Contig128 (25935-26387 m)	60	<p>Identities = 58/144 (40%), Positives = 92/144 (63%) gb AAK05931.1 AE006413_7 (AE006413) conserved hypothetical protein [Lactococcus lactis subsp. lactis] Length = 145</p>
SeqID 685	SA-1586.2	Contig128 (26405-27607 m)	48	<p>Identities = 106/401 (26%), Positives = 199/401 (49%), Gaps = 21/401 (5%) pir B70065 antibiotic resistance protein homolog ywoG - Bacillus subtilis emb CAB05383.1 (Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis] emb CAB15662.1 (Z99122) similar to antibiotic resistance protein [Bacillus subtilis] Length = 396</p>
SeqID 686	SA-1587.2	Contig139 (97176-98552 p)	78	<p>Identities = 280/458 (61%), Positives = 363/458 (79%), Gaps = 3/458 (0%) gb AAD47594.1 AF140784_3 (AF140784) Vexp3 [Streptococcus pneumoniae] Length = 459</p>
SeqID 687	SA-1588.1	Contig139 (96520-97176 p)	76	<p>Identities = 142/207 (68%), Positives = 169/207 (81%) gb AAD47593.1 AF140784_2 (AF140784) Vexp2 [Streptococcus pneumoniae] Length = 215</p>
SeqID 688	SA-1589.1	Contig139 (95233-96510 p)	63	<p>Identities = 165/425 (38%), Positives = 271/425 (62%), Gaps = 4/425 (0%) gb AAD47592.1 AF140784_1 (AF140784) Vexp1 [Streptococcus pneumoniae] Length = 425</p>
SeqID 689	SA-159.1	Contig138 (77622-78329 m)	32	<p>Identities = 44/150 (29%), Positives = 80/150 (53%), Gaps = 14/150 (9%) sp P16396 SUBE_BACSU MINOR EXTRACELLULAR PROTEASE EPR PRECURSOR pir SUBSMP serine proteinase (EC 3.4.21.-) epr precursor - Bacillus subtilis emb CAA37392.1 (X53307) prepropeptide (AA - 27 to 618) [Bacillus subtilis] gb AAA22423.1 (M22407) extracellular protease precursor [Bacillus subtilis] emb CAA51571.1 (X73124) ipa-15r epr [Bacillus subtilis] emb CAB15866.1 (Z99123) extracellular serine protease [Bacillus subtilis] Length = 645</p>
SeqID 690	SA-1590.1	Contig139 (94330-94491 m)	No Hits found	

SeqID 691	SA-1591.1	Contig139 (94053-94157 p)	65	<p>Identities = 40/88 (45%), Positives = 60/88 (67%) ref NP_065294.1 100 pct identical to sp:Y15B_EC0L[hypothetical 33.3 kd protein (orfB) of IS150] [Escherichia coli] pir H65154 probable transposase, 33.3K - Escherichia coli insertion sequence IS150 gb AAC76582.1 (AE000433) IS150 putative transposase [Escherichia coli K12] dbj BAB12587.1 (AP002527) 100 pct identical to sp:Y15B_EC0L[hypothetical 33.3 kd protein (orfB) of IS150] [Escherichia coli] Length = 283</p>
SeqID 692	SA-1593.1	Contig139 (93562-93822 p)	50	<p>Identities = 33/76 (43%), Positives = 44/76 (57%), Gaps = 1/76 (1%) ref NP_052792.1 pXO1-96 [Bacillus anthracis] pir H59102 hypothetical protein pXO1-96 - Bacillus anthracis virulence plasmid pXO1 gb AAA74027.1 (U30714) ORFB [Bacillus anthracis] gb AAA74029.1 (U30715) ORFB [Bacillus anthracis] gb AAD32400.1 AAD32400 (AF065404) pXO1-96 [Bacillus anthracis] Length = 274</p>
SeqID 693	SA-1594.1	Contig139 (93367-93540 p)	37	<p>Identities = 29/59 (49%), Positives = 39/59 (65%) gb AAB00677.1 (L40841) transposase [Enterococcus faecium] Length = 310</p>
SeqID 694	SA-1595.1	Contig139 (92774-93178 p)	36	<p>Identities = 28/116 (24%), Positives = 59/116 (50%), Gaps = 6/116 (5%) sp Q48585 Y13A_LACJO INSERTION ELEMENT IS1223 HYPOTHETICAL 20.7 KD PROTEIN (ORFA) gb AAA56999.1 (U09558) ORFA, putative Helix-Turn-Helix motif from amino acid 21 through 42 and from amino acid 78 through 99 [Lactobacillus johnsonii] Length = 177</p>
SeqID 695	SA-1596.1	Contig139 (91891-92511 m)	67	<p>Identities = 115/194 (59%), Positives = 139/194 (71%), Gaps = 11/194 (5%) pir T13289 probable integrase - Streptococcus phage phi-O1205 gb AAC79517.1 (U88974) ORF1 [Streptococcus thermophilus temperate bacteriophage O1205] Length = 359</p>

SeqID 696	SA-1597.1	Contig139 (91399-91746 p)	78	Identities = 110/115 (95%), Positives = 112/115 (96%) sp O34031 RL19_STRTR_50S_RIBOSOMAL_PROTEIN_L19 gb AAC01534.1 (U88973) ribosomal protein L19 [Streptococcus thermophilus] Length = 115
SeqID 697	SA-1598.1	Contig139 (89623-90819 p)	51	Identities = 131/393 (33%), Positives = 206/393 (52%), Gaps = 20/393 (5%) pir T44296 hypothetical protein [imported] - Bacillus halodurans dbj BAA75315.1 (AB011836) similar to Bordetella paraperlussis transposase for insertion sequence element(27 -identity) [Bacillus halodurans] dbj BAB04382.1 (AP001509) unknown conserved protein in others [Bacillus halodurans] Length = 424
SeqID 698	SA-1599.1	Contig139 (89355-89630 p)	52	Identities = 24/89 (26%), Positives = 50/89 (55%), Gaps = 2/89 (2%) pir G71950 hypothetical protein jhp0276 - Helicobacter pylori (strain J99) gb AAD05867.1 (AE001465) putative [Helicobacter pylori J99] Length = 100
SeqID 699	SA-16.1	Contig137 (28276-28605 m)	41	Identities = 35/103 (33%), Positives = 46/103 (43%), Gaps = 14/103 (13%) gb AAC02237.1 (U72957) merozoite surface protein 2 [Plasmodium falciparum] Length = 176
SeqID 700	SA-160.2	Contig138 (76196-77368 m)	71	Identities = 236/387 (60%), Positives = 281/387 (71%), Gaps = 10/387 (2%) gb AAK05433.1 AE006366_2 (AE006366) N- acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) [Lactococcus lactis subsp. lactis] Length = 382
SeqID 701	SA-1601.1	Contig139 (88835-89278 p)	55	Identities = 63/146 (43%), Positives = 87/146 (59%), Gaps = 1/146 (0%) sp O34589 FLAW_BACSU_PROBABLE FLAVODOXIN 2 pir E69866 flavodoxin homolog ykuP[similarity] Bacillus subtilis emb CAA10879.1 (AJ222587) YkuP protein [Bacillus subtilis] emb CAB13290.1 (Z99111) similar to sulfite reductase [Bacillus subtilis] Length = 178

SeqID 702	SA-1602.1	Contig139 (87754-88776 p)	51	Identities = 123/343 (35%), Positives = 175/343 (50%), Gaps = 26/343 (7%) pir T35340 probable adenosine deaminase - Streptomyces coelicolor emb CAB42949.1 (AL049863) putative adenosine deaminase [Streptomyces coelicolor A3(2)] Length = 359
SeqID 703	SA-1603.2	Contig139 (86553-87488 m)	66	Identities = 144/306 (47%), Positives = 209/306 (68%), Gaps = 4/306 (1%) pir F69999 conserved hypothetical protein ytl - Bacillus subtilis gb AAC00337.1 (AF008220) Ytl [Bacillus subtilis] emb CAB14885.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 313
SeqID 704	SA-1604.1	Contig90 (12596-12865 m)	76	Identities = 67/89 (75%), Positives = 76/89 (85%) sp O31587 R14B_BACSU 30S RIBOSOMAL PROTEIN S14-2 pir F69835 ribosomal protein S14 homolog yhzA - Bacillus subtilis emb CAB12716.1 (Z99108) similar to ribosomal protein S14 [Bacillus subtilis] Length = 89
SeqID 705	SA-1605.2	Contig90 (11209-12234 p)	58	Identities = 139/339 (41%), Positives = 201/339 (59%), Gaps = 4/339 (1%) ref NP_047071.1 L4171.5 [Leishmania major] pir T02833 threonine aldolase L4171.5 [imported] - Leishmania major (strain Friedlin) gb AAC24656.1 (AE001274) L4171.5 [Leishmania major] Length = 359
SeqID 706	SA-1606.2	Contig90 (10079-11089 p)	81	Identities = 244/343 (71%), Positives = 280/343 (81%), Gaps = 7/343 (2%) gb AAK04393.1 AE006266_12 (AE006266) O-sialoglycoprotein endopeptidase [Lactococcus lactis subsp. lactis] Length = 346
SeqID 707	SA-1607.1	Contig90 (9557-10003 p)	48	Identities = 44/141 (31%), Positives = 73/141 (51%), Gaps = 8/141 (5%) pir B70351 ribosomal-protein-alanine acetyltransferase - Aquifex aeolicus gb AAC06803.1 (AE000696) ribosomal-protein-alanine acetyltransferase [Aquifex aeolicus] Length = 154

SeqID 708	SA-1608.1	Contig90 (8905-9594 p)	54	Identities = 94/182 (51%), Positives = 127/182 (69%), Gaps = 6/182 (3%) emb CAA76861.1 (Y17797) hypothetical protein [Enterococcus faecalis] Length = 204
SeqID 709	SA-1609.1	Contig90 (8493-8723 m)	55	Identities = 30/67 (44%), Positives = 45/67 (66%), Gaps = 7/67 (10%) dbj BAB06380.1 (AP001516) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 69
SeqID 710	SA-161.2	Contig136 (90788-91228 m)	47	Identities = 37/142 (26%), Positives = 71/142 (49%), Gaps = 7/142 (4%) pir B70057 transcription regulator MarR family homolog ywhA - Bacillus subtilis emb CAB02511.1 (Z80360) Unknown, similar to regulatory proteins [Bacillus subtilis] emb CAB15782.1 (Z99123) similar to transcriptional regulator (MarR family) [Bacillus subtilis] Length = 139
SeqID 711	SA-1610.1	Contig90 (6760-8439 m)	77	Identities = 351/550 (63%), Positives = 442/550 (79%) dbj BAB06381.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 555
SeqID 712	SA-1611.1	Contig90 (6089-6598 p)	38	Identities = 45/119 (37%), Positives = 67/119 (55%), Gaps = 4/119 (3%) emb CAA66624.1 (X97985) ORF1 [Staphylococcus aureus] Length = 255
SeqID 713	SA-1612.1	Contig90 (4595-5941 p)	89	Identities = 390/449 (86%), Positives = 406/449 (89%), Gaps = 1/449 (0%) gb AAC44800.1 (U61271) glutamine synthetase type 1 [Streptococcus agalactiae] Length = 449
SeqID 714	SA-1613.1	Contig90 (4190-4561 p)	69	Identities = 59/123 (47%), Positives = 89/123 (71%), Gaps = 5/123 (4%) sp P19083 GLNR_BACCE REGULATORY PROTEIN GLNR pir JU0076 hypothetical 15K protein (glnA 5 region) - Bacillus cereus dbj BAA00402.1 (D00513) ORF129 [Bacillus cereus] Length = 129
SeqID 715	SA-1614.1	Contig90 (3628-4110 p)	33	Identities = 35/146 (23%), Positives = 64/146 (42%), Gaps = 7/146 (4%) dbj BAB04661.1 (AP001510) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 360

SeqID 716	SA-1615.1	Contig90 (2945-3355 m)	56	Identities = 61/134 (45%), Positives = 78/134 (57%), Gaps = 5/134 (3%) ref NP_00939.1 Ycr013cp [Saccharomyces cerevisiae] sp P25614 YCQ3_YEAST VERY HYPOTHETICAL 22.8 KD PROTEIN IN PGK1-POL4 INTERGENIC REGION pir S19423 probable membrane protein YCR013c - yeast (Saccharomyces cerevisiae) emb CAA42330.1 (X59720) YCR013c, len:215 [Saccharomyces cerevisiae] Length = 215
SeqID 717	SA-1616.1	Contig90 (2112-3308 p)	75	Identities = 254/398 (63%), Positives = 303/398 (75%), Gaps = 2/398 (0%) sp Q9Z5C4 PGK_STAAU PHOSPHOGLYCERATE KINASE emb CAB38646.1 (AJ133520) phosphoglycerate kinase [Staphylococcus aureus] Length = 396
SeqID 718	SA-1617.2	Contig90 (1108-1977 p)	81	Identities = 182/291 (62%), Positives = 237/291 (80%), Gaps = 7/291 (2%) emb CAA73175.1 (Y12602) acid phosphatase [Streptococcus equisimilis] Length = 285
SeqID 719	SA-1619.1	Contig108 (18754-19293 m)	62	Identities = 89/146 (60%), Positives = 113/146 (76%), Gaps = 1/146 (0%) gb AAAF08325.1 U78969_2 (U78969) FlaR [Streptococcus pyogenes] Length = 172
SeqID 720	SA-162.1	Contig136 (88924-90741 m)	52	Identities = 216/543 (39%), Positives = 322/543 (58%), Gaps = 25/543 (4%) emb CAB69751.1 (AL137187) putative ABC transporter [Streptomyces coelicolor A3(2)] Length = 577
SeqID 721	SA-1620.1	Contig108 (17468-18553 m)	43	Identities = 94/307 (30%), Positives = 157/307 (50%), Gaps = 21/307 (6%) pir G72378 glycerol dehydrogenase - Thermotoga maritima (strain MSB8) gb AAD35508.1 AE001721_7 (AE001721) glycerol dehydrogenase [Thermotoga maritima] Length = 364
SeqID 722	SA-1621.1	Contig108 (16596-17303 p)	23	Identities = 38/98 (38%), Positives = 59/98 (59%) emb CAC01354.1 (AL390975) putative integral membrane protein [Streptomyces coelicolor A3(2)] Length = 198

SeqID 723	SA-1622.1	Contig108 (14405-16249 p)	69	Identities = 301/610 (49%), Positives = 437/610 (71%), Gaps = 9/610 (1%) dbj BAB05348.1 (AP001512) unknown conserved protein [Bacillus halodurans] Length = 618
SeqID 724	SA-1623.1	Contig108 (12108-14360 p)	68	Identities = 397/751 (52%), Positives = 519/751 (68%), Gaps = 14/751 (1%) dbj BAB04157.1 (AP001508) homoserine methyl transferase [Bacillus halodurans] Length = 756
SeqID 725	SA-1624.1	Contig108 (11622-11753 p)	No Hits found	
SeqID 726	SA-1625.2	Contig108 (10745-11440 p)	60	Identities = 86/225 (38%), Positives = 142/225 (62%), Gaps = 1/225 (0%) gb AAK05711.1 AE006391_10 (AE006391) amino acid permease [Lactococcus lactis subsp. lactis] Length = 235
SeqID 727	SA-1626.1	Contig87 (8382-8996 m)	No Hits found	
SeqID 728	SA-1627.1	Contig87 (7700-8284 m)	No Hits found	
SeqID 729	SA-1628.1	Contig87 (7137-7703 m)	No Hits found	
SeqID 730	SA-1629.1	Contig87 (4483-7137 m)	74	Identities = 544/880 (61%), Positives = 676/880 (76%), Gaps = 11/880 (1%) pir SYBSVS valine-tRNA ligase (EC 6.1.1.9) - Bacillus stearothermophilus Length = 880
SeqID 731	SA-163.1	Contig136 (87180-88934 m)	55	Identities = 228/526 (43%), Positives = 331/526 (62%), Gaps = 2/526 (0%) emb CAB69752.1 (AL137187) putative ABC transporter [Streptomyces coelicolor A3(2)] Length = 642
SeqID 732	SA-1630.1	Contig87 (3318-4247 p)	34	Identities = 57/231 (24%), Positives = 109/231 (46%), Gaps = 10/231 (4%) pir T35598 hypothetical protein SC6G9.01c - Streptomyces coelicolor (fragment) emb CAB45592.1 (AL079356) hypothetical protein [Streptomyces coelicolor] Length = 409
SeqID 733	SA-1632.1	Contig87 (1942-2862 m)	43	Identities = 100/306 (32%), Positives = 152/306 (48%), Gaps = 3/306 (0%) pir T50594 probable oxidoreductase [imported] - Streptomyces coelicolor emb CAB61731.1 (AL133220) putative oxidoreductase. [Streptomyces coelicolor A3(2)] Length = 351

SeqID 734	SA-1634.1	Contig87 (879-1781 m)	44	Identities = 70/284 (24%), Positives = 135/284 (46%), Gaps = 8/284 (2%) ref NP_053049.1 hypothetical protein [Plasmid pNZ4000] gb AAD40365.1 (AF036485) hypothetical protein [Plasmid pNZ4000] Length = 302
SeqID 735	SA-1635.1	Contig87 (3-611 m)	23	Identities = 38/107 (35%), Positives = 59/107 (54%), Gaps = 6/107 (5%) sp P37507 YQAQ_BACSU HYPOTHETICAL 13.9 KD PROTEIN IN COTF-TETB INTERGENIC REGION pir S66000 yyaQ protein - Bacillus subtilis dbj BAA05206.1 (D26185) unknown [Bacillus subtilis] emb CAB16112.1 (Z99124) yyaQ [Bacillus subtilis] Length = 118
SeqID 736	SA-1636.2	Contig124 (24911-26119 m)	73	Identities = 234/384 (60%), Positives = 300/384 (77%), Gaps = 2/384 (0%) sp P42602 YGJU_ECOLI HYPOTHETICAL SYMPORTER YGJU pir F65097 hypothetical 43.5 kD protein in ebgC-exuT intergenic region - Escherichia coli (strain K-12) gb AAA57891.1 (U18997) ORF_o414; Geneplot suggests frameshift near start but none found [Escherichia coli] gb AAC76124.1 (AE000391) putative transport protein [Escherichia coli K12] Length = 414
SeqID 737	SA-1638.1	Contig124 (24194-24742 m)	61	Identities = 109/182 (59%), Positives = 141/182 (76%) emb CAB59830.1 (AJ012388) hypothetical protein [Lactococcus lactis] Length = 182
SeqID 738	SA-164.1	Contig136 (86446-87072 m)	51	Identities = 67/182 (36%), Positives = 111/182 (60%), Gaps = 4/182 (2%) dbj BAB05604.1 (AP001513) unknown conserved protein [Bacillus halodurans] Length = 251
SeqID 739	SA-1640.2	Contig124 (22501-24177 m)	44	Identities = 150/481 (31%), Positives = 247/481 (51%), Gaps = 15/481 (3%) pir H69858 cation ABC transporter (ATP-binding protein) homolog ykoD - Bacillus subtilis emb CAA05601.1 (AJ002571) YkoD [Bacillus subtilis] emb CAB13179.1 (Z99110) similar to cation ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 490

SeqID 740	SA-1641.2	Contig124 (21678-22508 m)	25	Identities = 46/134 (34%), Positives = 74/134 (54%), Gaps = 3/134 (2%) pir F75200 hypothetical protein PAB2261 - Pyrococcus abyssi (strain Orsay) emb CAB49053.1 (AJ248283) hypothetical protein [Pyrococcus abyssi] Length = 248
SeqID 741	SA-1642.1	Contig124 (20972-21643 p)	49	Identities = 70/183 (38%), Positives = 113/183 (61%), Gaps = 2/183 (1%) gb AAC46144.1 (AF001974) putative TrkA [Thermoanaerobacter ethanolicus] Length = 195
SeqID 742	SA-1643.1	Contig124 (19568-20959 p)	55	Identities = 170/461 (36%), Positives = 263/461 (56%), Gaps = 28/461 (6%) pir G53610 ntpJ protein - Enterococcus hirae Length = 448
SeqID 743	SA-1644.1	Contig124 (18852-19565 p)	70	Identities = 132/240 (55%), Positives = 171/240 (71%), Gaps = 4/240 (1%) sp P25813 GIDB_BACSU GLUCOSE INHIBITED DIVISION PROTEIN B pir BWBSGB glucose-inhibited division protein gidB - Bacillus subtilis emb CAA44405.1 (X62539) homologous to E.coli gidB [Bacillus subtilis] dbj BAA05230.1 (D26185) unknown [Bacillus subtilis] emb CAB16137.1 (Z99124) glucose-inhibited division protein [Bacillus subtilis] Length = 239
SeqID 744	SA-1645.1	Contig124 (18170-18724 m)	78	Identities = 124/182 (68%), Positives = 152/182 (83%) gb AAG23700.1 (AF017421) LemA-like protein [Streptococcus gordonii] Length = 189
SeqID 745	SA-1646.1	Contig124 (17259-18149 m)	85	Identities = 220/297 (74%), Positives = 261/297 (87%), Gaps = 1/297 (0%) sp O30795 HTPX_STRGC PROBABLE PROTEASE HTPX HOMOLOG pir T48855 probable heat shock protein HtpX [imported] - Streptococcus gordonii gb AAB70525.1 (AF017421) putative heat shock protein HtpX [Streptococcus gordonii] Length = 297
SeqID 746	SA-1647.1	Contig124 (16621-17154 m)	70	Identities = 92/174 (52%), Positives = 133/174 (75%), Gaps = 3/174 (1%) gb AAG32547.1 (U12643) YlbN-like hypothetical protein [Streptococcus gordonii] Length = 176

SeqID 747	SA-1648.1	Contig124 (15697-16386 m)	91	Identities = 193/229 (84%), Positives = 211/229 (91%), Gaps = 1/229 (0%) gb AAC64935.1 (AF082668) CsrR [Streptococcus pyogenes] Length = 228
SeqID 748	SA-1649.2	Contig124 (14202-15707 m)	72	Identities = 248/501 (49%), Positives = 364/501 (72%), Gaps = 4/501 (0%) gb AAC64936.1 (AF082668) CsrS [Streptococcus pyogenes] Length = 500
SeqID 749	SA-165.1	Contig136 (85758-86465 m)	58	Identities = 84/226 (37%), Positives = 138/226 (60%), Gaps = 9/226 (3%) dbj BAB04091.1 (AP001508) two-component response regulator [Bacillus halodurans] Length = 222
SeqID 750	SA-1650.2	Contig134 (63460-63936 p)	41	Identities = 49/136 (36%), Positives = 69/136 (50%), Gaps = 8/136 (5%) pir G75571 MutT/nudix family protein - Deinococcus radiodurans (strain R1) gb AAF09597.1 AE001864_3 (AE001864) MutT/nudix family protein [Deinococcus radiodurans] Length = 350
SeqID 751	SA-1651.1	Contig134 (64258-65064 p)	47	Identities = 70/254 (27%), Positives = 127/254 (49%), Gaps = 2/254 (0%) pir G83412 hypothetical protein PA1860 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG05249.1 AE004612_3 (AE004612) hypothetical protein [Pseudomonas aeruginosa] Length = 275
SeqID 752	SA-1653.1	Contig134 (65459-65956 p)	84	Identities = 140/164 (85%), Positives = 157/164 (95%) gb AAC38046.1 (AF000954) No definition line found [Streptococcus mutans] Length = 164
SeqID 753	SA-1654.1	Contig134 (65937-66335 p)	90	Identities = 109/133 (81%), Positives = 122/133 (90%), Gaps = 2/133 (1%) pir A36933 diacylglycerol kinase homolog - Streptococcus mutans Length = 137
SeqID 754	SA-1655.1	Contig134 (66332-67276 p)	83	Identities = 273/299 (91%), Positives = 290/299 (96%) gb AAD41632.1 AF072811_1 (AF072811) GTPase Era [Streptococcus pneumoniae] Length = 299
SeqID 755	SA-1656.1	Contig134 (67495-67914 p)	No Hits found	
SeqID 756	SA-1657.1	Contig134 (68290-68625 p)	No Hits found	

SeqID 757	SA-1658.1	Contig134 (68686-68898 p)	52	Identities = 31/87 (35%), Positives = 48/87 (54%), Gaps = 1/87 (1%) emb CAC03528.1 (AJ276410) BlpY protein [Streptococcus pneumoniae] Length = 229
SeqID 758	SA-1659.1	Contig134 (69367-69447 p)	No Hits found	
SeqID 759	SA-166.1	Contig136 (84529-85758 m)	35	Identities = 76/315 (24%), Positives = 149/315 (47%), Gaps = 33/315 (10%) gb AAF79919.1 (AF039082) putative histidine protein kinase [Lactococcus lactis] Length = 414
SeqID 760	SA-1660.1	Contig134 (69841-71433 m)	10	Identities = 36/119 (30%), Positives = 57/119 (47%), Gaps = 15/119 (12%) pir S74641 hypothetical protein sll1681 - Synechocystis sp. (strain PCC 6803) dbj BAA16793.1 (D90900) hypothetical protein [Synechocystis sp.] Length = 507
SeqID 761	SA-1662.1	Contig134 (71635-72498 p)	48	Identities = 71/273 (26%), Positives = 140/273 (51%), Gaps = 16/273 (5%) sp P49330 RGG_STRGC RGG PROTEIN pir A41898 positive regulator of glucosyltransferase and Spp phenotype rgg - Streptococcus gordonii gb AAA26968.1 (M89776) rgg [Streptococcus gordonii] Length = 297
SeqID 762	SA-1663.1	Contig134 (72647-73468 p)	77	Identities = 182/271 (67%), Positives = 217/271 (79%) sp P55045 FPG_STRMU FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE) dbj BAA05066.1 (D26071) formamidopyrimidine-DNA glycosylase [Streptococcus mutans] Length = 273
SeqID 763	SA-1665.1	Contig134 (73444-74052 p)	52	Identities = 80/189 (42%), Positives = 113/189 (59%), Gaps = 1/189 (0%) sp O34932 YTAG_BACSU HYPOTHETICAL 22.0 KD PROTEIN IN GAPB-MUTM INTERGENIC REGION pir A69988 conserved hypothetical protein ytaG - Bacillus subtilis gb AAC00353.1 (AF008220) YtaG [Bacillus subtilis] emb CAB14866.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 197

SeqID 764	SA-1667.1	Contig134 (74176-74877 p)	51	Identities = 68/247 (27%), Positives = 126/247 (50%), Gaps = 17/247 (6%) pir E69771 ABC transporter (ATP-binding protein) homolog ydbJ - Bacillus subtilis dbj BAA19286.1 (AB001488) PROBABLE TRANSPORT ATP BINDING PROTEIN. [Bacillus subtilis] emb CAB12256.1 (Z99106) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 308
SeqID 765	SA-1669.1	Contig134 (74855-76513 p)	No Hits found	
SeqID 766	SA-167.1	Contig136 (83535-84413 m)	74	Identities = 183/290 (63%), Positives = 223/290 (76%) gb AAG02450.1 AF290097_1 (AF290097) mevalonate kinase [Streptococcus pyogenes] Length = 292
SeqID 767	SA-1671.2	Contig134 (76757-77878 p)	77	Identities = 230/396 (58%), Positives = 315/396 (79%) emb CAA07482.1 (AJ007367) multi-drug resistance efflux pump [Streptococcus pneumoniae] Length = 399
SeqID 768	SA-1673.2	Contig133 (6150-7757 m)	87	Identities = 413/535 (77%), Positives = 476/535 (88%), Gaps = 1/535 (0%) sp Q59905 DEXB_STREQ GLUCAN 1,6-ALPHA-GLUCOSIDASE (DEXTRAN GLUCOSIDASE) (EXO-1,6-ALPHA-GLUCOSIDASE) (GLUCODEXTRANASE) pir S39970 glucan 1,6-alpha-glucosidase (EC 3.2.1.70) - Streptococcus equisimilis emb CAA51348.1 (X72832) glucan 1,6-alpha-glucosidase [Streptococcus equisimilis] prf 2009358A glucan 1,6-alpha-glucosidase [Streptococcus equisimilis] Length = 537
SeqID 769	SA-1674.1	Contig133 (5067-6062 m)	91	Identities = 267/331 (80%), Positives = 306/331 (91%) sp P96995 GALE_STRMU UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) pir JC5313 UDPglucose 4-epimerase (EC 5.1.3.2) - Streptococcus mutans gb AAB49738.1 (U21942) UDP-galactose 4-epimerase [Streptococcus mutans] Length = 333

SeqID 770	SA-1675.1	Contig133 (4366-5046 p)	58	Identities = 71/223 (31%), Positives = 139/223 (61%), Gaps = 7/223 (3%) dbj BAB06470.1 (AP001516) two-component response regulator [Bacillus halodurans] Length = 230
SeqID 771	SA-1676.1	Contig133 (2946-4364 p)	47	Identities = 132/461 (28%), Positives = 245/461 (52%), Gaps = 7/461 (1%) pir A70009 two-component sensor histidine kinase homolog yufL - Bacillus subtilis emb CAB07946.1 (Z93937) unknown [Bacillus subtilis] emb CAB15141.1 (Z99120) similar to two-component sensor histidine kinase [YufM] [Bacillus subtilis] Length = 533
SeqID 772	SA-1677.1	Contig133 (1347-2684 m)	82	Identities = 329/428 (76%), Positives = 375/428 (86%) gb AAB18291.1 (U35658) L-malate permease [Streptococcus bovis] Length = 441
SeqID 773	SA-1678.2	Contig133 (628-1212 p)	22	Identities = 36/70 (51%), Positives = 44/70 (62%) pir B72732 hypothetical protein APE0395 - Aeropyrum pernix (strain K1) dbj BAA79350.1 (AP000059) 127aa long hypothetical protein [Aeropyrum pernix] Length = 127
SeqID 774	SA-1679.2	Contig133 (159-1322 m)	83	Identities = 285/386 (73%), Positives = 332/386 (85%), Gaps = 1/386 (0%) gb AAB07709.1 (U35659) malic enzyme [Streptococcus bovis] Length = 389
SeqID 775	SA-168.1	Contig136 (82609-83553 m)	80	Identities = 220/312 (70%), Positives = 264/312 (84%) gb AAG02456.1 AF290099_2 (AF290099) mevalonate diphosphate decarboxylase [Streptococcus pneumoniae] Length = 317
SeqID 776	SA-1680.2	Contig96 (2340-2804 p)	59	Identities = 74/146 (50%), Positives = 103/146 (69%), Gaps = 3/146 (2%) gb AAK04739.1 AE006297_2 (AE006297) transcriptional regulator CtsR [Lactococcus lactis subsp. lactis] Length = 151
SeqID 777	SA-1681.2	Contig96 (2801-5248 p)	69	Identities = 404/831 (48%), Positives = 567/831 (67%), Gaps = 52/831 (6%) gb AAK04740.1 AE006297_3 (AE006297) ATP-dependent protease ATP-binding subunit [Lactococcus lactis subsp. lactis] Length = 816

SeqID 778	SA-1682.1	Contig96 (5468-5959 m)	50	Identities = 53/156 (33%), Positives = 84/156 (52%) gb AAF15587.1 AF187951_1 (AF187951) promotes resistance to glutamine synthetase inhibitors [Activation-tagging vector pSKI015] Length = 183
SeqID 779	SA-1683.1	Contig96 (5973-6614 m)	85	Identities = 179/213 (84%), Positives = 197/213 (92%) gb AAC97156.1 (U49397) unknown [Streptococcus pyogenes] Length = 213
SeqID 780	SA-1685.1	Contig96 (6735-7712 m)	93	Identities = 282/325 (86%), Positives = 306/325 (93%) gb AAC97155.1 (U49397) unknown [Streptococcus pyogenes] Length = 325
SeqID 781	SA-1686.1	Contig96 (7696-8571 m)	84	Identities = 224/271 (82%), Positives = 247/271 (90%) sp Q9ZB45 HSLO_STRPY_33 KDA CHAPERONIN (HEAT SHOCK PROTEIN_33 HOMOLOG) (HSP33) gb AAC97154.1 (U49397) unknown [Streptococcus pyogenes] Length = 277
SeqID 782	SA-1687.1	Contig96 (8711-9967 m)	34	Identities = 64/350 (18%), Positives = 146/350 (41%), Gaps = 7/350 (2%) gb AAB93480.1 (AF019377) tellurite resistance protein [Rhodobacter sphaeroides] Length = 396
SeqID 783	SA-1689.2	Contig96 (9964-10782 m)	37	Identities = 56/238 (23%), Positives = 105/238 (43%), Gaps = 25/238 (10%) emb CAB39037.2 (AL034559) hypothetical protein, PFC0960c [Plasmodium falciparum] Length = 1946
SeqID 784	SA-169.1	Contig136 (81624-82616 m)	70	Identities = 170/330 (51%), Positives = 233/330 (70%), Gaps = 1/330 (0%) gb AAG02457.1 AF290099_3 (AF290099) phosphomevalonate kinase [Streptococcus pneumoniae] Length = 336
SeqID 785	SA-1690.2	Contig96 (10906-11172 m)	55	Identities = 35/76 (46%), Positives = 49/76 (64%) gb AAK03132.1 (AE006146) unknown [Pasteurella multocida] Length = 757
SeqID 786	SA-1694.1	Contig83 (3183-4526 m)	37	Identities = 143/211 (67%), Positives = 170/211 (79%), Gaps = 3/211 (1%) gb AAD00288.1 (U78607) putative secreted protein [Streptococcus mutans] Length = 211
SeqID 787	SA-1695.1	Contig83 (2091-3071 m)	78	Identities = 209/312 (66%), Positives = 266/312 (84%), Gaps = 3/312 (0%) gb AAA62181.1 (M92842) prs [Listeria monocytogenes] Length = 318

SeqID 788	SA-1696.1	Contig83 (808-1986 m)	74	Identities = 261/391 (66%), Positives = 323/391 (81%) gb AAF06954.1 AF146529_1 (AF146529) aromatic amino acid aminotransferase [Lactococcus lactis subsp. cremoris] Length = 391
SeqID 789	SA-1697.1	Contig83 (57-818 m)	47	Identities = 81/242 (33%), Positives = 124/242 (50%), Gaps = 3/242 (1%) dbj BAB05088.1 (AP001511) unknown conserved protein [Bacillus halodurans] Length = 254
SeqID 790	SA-1698.2	Contig124 (4176-4679 m)	49	Identities = 70/150 (46%), Positives = 100/150 (66%), Gaps = 2/150 (1%) gb AAK06154.1 AE006436_3 (AE006436) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 164
SeqID 791	SA-1699.1	Contig124 (4765-6096 m)	70	Identities = 237/432 (54%), Positives = 315/432 (72%), Gaps = 3/432 (0%) sp P40778 MURC_BACSU UDP-N- ACETYLMURAMATE--ALANINE LIGASE (UDP-N- ACETYLMURAMOYL-L-ALANINE SYNTHETASE) pir C69662 UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8) murC - Bacillus subtilis gb AAC00294.1 (AF008220) putative UDP-N-acetylmuramate-alanine ligase [Bacillus subtilis] emb CAB14957.1 (Z99119) UDP-N-acetyl muramate-alanine ligase [Bacillus subtilis] Length = 432
SeqID 792	SA-17.1	Contig137 (28295-29017 p)	49	Identities = 82/270 (30%), Positives = 124/270 (45%), Gaps = 34/270 (12%) pir S45091 hypothetical protein iota - Streptococcus pyogenes plasmid pDB101 pir S68125 hypothetical protein iota - Streptococcus pyogenes plasmid pSM19035 emb CAA47097.1 (X66468) orf iota [Streptococcus pyogenes] Length = 268
SeqID 793	SA-170.1	Contig136 (80632-81627 m)	58	Identities = 127/331 (38%), Positives = 195/331 (58%), Gaps = 9/331 (2%) dbj BAB07793.1 (AB037666) hypothetical protein [Streptomyces sp. CL190] Length = 363
SeqID 794	SA-1700.1	Contig124 (6106-6696 m)	No Hits found	
SeqID 795	SA-1702.1	Contig124 (6888-9986 m)	38	Identities = 260/678 (38%), Positives = 405/678 (59%), Gaps = 21/678 (3%) emb CAA67095.1 (X98455) SNF [Bacillus cereus] Length = 1064

SeqID 796	SA-1704.1	Contig124 (10142-11452 m)	92	Identities = 377/436 (86%), Positives = 414/436 (94%) dbj BAA88823.1 (AB016077) phosphoglycerate dehydrogenase [Streptococcus mutans] Length = 436
SeqID 797	SA-1705.1	Contig124 (11500-12402 m)	63	Identities = 105/313 (33%), Positives = 191/313 (60%), Gaps = 17/313 (5%) sp P06567 DNA_BACSU PRIMOSOMAL PROTEIN DNAI pir IQBS44 primosome component (helicase loader) dnaI - Bacillus subtilis emb CAA99605.1 (X04963) ORF 311 (AA 1- 311) [Bacillus subtilis] emb AAC00359.1 (AF008220) DnaI protein [Bacillus subtilis] gb CAB14858.1 (Z99118) helicase loader [Bacillus subtilis] Length = 311
SeqID 798	SA-1706.2	Contig124 (12399-13574 m)	52	Identities = 125/388 (32%), Positives = 207/388 (53%), Gaps = 14/388 (3%) gb AAK04849.1 AE006308_9 (AE006308) replication protein DnaB [Lactococcus lactis subsp. lactis] Length = 391
SeqID 799	SA-1707.2	Contig124 (13574-14053 m)	72	Identities = 100/150 (66%), Positives = 119/150 (78%), Gaps = 2/150 (1%) gb AAK04848.1 AE006308_8 (AE006308) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 148
SeqID 800	SA-171.1	Contig136 (80346-80564 m)	No Hits found	
SeqID 801	SA-1710.1	Contig98 (11651-12985 m)	No Hits found	
SeqID 802	SA-1711.1	Contig98 (10742-11626 p)	42	Identities = 97/251 (38%), Positives = 147/251 (57%), Gaps = 15/251 (5%) sp P54604 YHCT_BACSU HYPOTHETICAL 33.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION pir H69823 conserved hypothetical protein yhcT - Bacillus subtilis emb CAA65704.1 (X96983) hypothetical protein [Bacillus subtilis] emb CAB12749.1 (Z99108) similar to hypothetical proteins [Bacillus subtilis] Length = 302
SeqID 803	SA-1712.1	Contig98 (8377-10698 m)	71	Identities = 424/773 (54%), Positives = 555/773 (70%), Gaps = 47/773 (6%) gb AAF04735.1 AF101780_1 (AF101780) penicillin- binding protein 2a [Streptococcus pneumoniae] Length = 731

SeqID 804	SA-1714.1	Contig98 (7967-8140 m)	No Hits found	
SeqID 805	SA-1715.2	Contig98 (5728-6195 p)	49	Identities = 54/136 (39%), Positives = 77/136 (55%), Gaps = 19/136 (13%) pir JQ0138 hypothetical 18.2K protein - Pseudomonas aeruginosa Length = 202
SeqID 806	SA-1716.2	Contig98 (5724-7655 m)	22	Identities = 96/302 (31%), Positives = 143/302 (46%), Gaps = 37/302 (12%) gb AAB71985.1 (U73163) M-like protein [Streptococcus equi] Length = 374
SeqID 807	SA-1718.2	Contig102 (7846-9225 p)	54	Identities = 147/473 (31%), Positives = 256/473 (54%), Gaps = 34/473 (7%) pir D69159 methyl coenzyme M reductase system, component A2 homolog - Methanobacterium thermoautotrophicum (strain Delta H) gb AAB84960.1 (AE000829) methyl coenzyme M reductase system, component A2 homolog [Methanothermobacter thermoautotrophicus] Length = 480
SeqID 808	SA-1719.1	Contig102 (7122-7853 p)	36	Identities = 49/192 (25%), Positives = 89/192 (45%), Gaps = 6/192 (3%) pir C69159 conserved hypothetical protein MTH453 - Methanobacterium thermoautotrophicum (strain Delta H) gb AAB84959.1 (AE000829) conserved protein [Methanobacterium thermoautotrophicum] Length = 252
SeqID 809	SA-172.1	Contig136 (79446-80300 p)	48	Identities = 86/291 (29%), Positives = 151/291 (51%), Gaps = 28/291 (9%) dbj BAB10885.1 (AB010693) gene_id:K21C13.21~pir T04769~strong similarity to unknown protein [Arabidopsis thaliana] Length = 325
SeqID 810	SA-1720.1	Contig102 (6544-7140 p)	No Hits found	
SeqID 811	SA-1722.1	Contig102 (6129-6533 p)	42	Identities = 38/153 (24%), Positives = 68/153 (43%), Gaps = 1/153 (0%) emb CAC17502.1 (AL450432) conserved hypothetical protein [Streptomyces coelicolor] Length = 677
SeqID 812	SA-1723.1	Contig102 (5905-6087 p)	No Hits found	

SeqID 813	SA-1724.1	Contig102 (5475-5918 p)	62	Identities = 77/221 (34%), Positives = 114/221 (50%), Gaps = 47/221 (21%) pir T30345 irp5 protein - Yersinia enterocolitica emb CAA73130.1 (Y12527) Irp5 protein [Yersinia enterocolitica] Length = 525
SeqID 814	SA-1725.1	Contig102 (4539-5375 p)	59	Identities = 111/270 (41%), Positives = 165/270 (61%), Gaps = 2/270 (0%) sp P40871 DHBE_BACSU 2,3-DIHYDROXYBENZOATE-AMP LIGASE (DIHYDROXYBENZOIC ACID-ACTIVATING ENZYME) pir D69615 2,3-dihydroxybenzoate--[carrier protein] ligase (EC 6.2.1.-) dhbE - Bacillus subtilis gb AAC44632.1 (U26444) 2,3-dihydroxybenzoate-AMP ligase [Bacillus subtilis] emb CAB15188.1 (Z99120) 2,3-dihydroxybenzoate-AMP ligase (enterobactin synthetase component E) [Bacillus subtilis] Length = 539
SeqID 815	SA-1726.1	Contig102 (3969-4343 p)	48	Identities = 43/113 (38%), Positives = 64/113 (56%), Gaps = 1/113 (0%) pir E71058 hypothetical protein PH1163 - Pyrococcus horikoshii dbj BAA30263.1 (AP000005) 148aa long hypothetical protein [Pyrococcus horikoshii] Length = 148
SeqID 816	SA-1728.1	Contig102 (3469-3921 p)	67	Identities = 78/151 (51%), Positives = 108/151 (70%) dbj BAB05053.1 (AP001511) late competence operon required for DNA binding and uptake [Bacillus halodurans] Length = 188
SeqID 817	SA-1729.1	Contig102 (2390-3457 p)	68	Identities = 183/352 (51%), Positives = 258/352 (72%), Gaps = 3/352 (0%) emb CAA70068.1 (Y08842) aminopeptidase P [Lactococcus lactis] Length = 352
SeqID 818	SA-173.1	Contig136 (78431-79345 p)	51	Identities = 83/308 (26%), Positives = 160/308 (51%), Gaps = 12/308 (3%) pir F69795 conserved hypothetical protein yerQ - Bacillus subtilis emb CAB12492.1 (Z99107) similar to hypothetical proteins [Bacillus subtilis] Length = 303
SeqID 819	SA-1730.1	Contig102 (1574-2290 p)	No Hits found	

SeqID 820	SA-1731.1	Contig102 (43-1572 p)	58	Identities = 198/481 (41%), Positives = 300/481 (62%), Gaps = 24/481 (4%) pir F69763 multidrug resistance protein homolog ycnB - Bacillus subtilis dbj BAA09016.1 (D50453) homologue of multidrug resistance protein B, EmrB, of E. coli [Bacillus subtilis] emb CAB12192.1 (Z99106) similar to multidrug resistance protein [Bacillus subtilis] Length = 472
SeqID 821	SA-1732.1	Contig130 (71572-71979 p)	67	Identities = 73/135 (54%), Positives = 94/135 (69%), Gaps = 5/135 (3%) gb AAB99856.1 (AF043542) nucleoside diphosphate kinase [Gallus gallus] Length = 153
SeqID 822	SA-1734.1	Contig130 (70987-71262 p)	No Hits found	
SeqID 823	SA-1736.1	Contig130 (69999-70358 p)	No Hits found	
SeqID 824	SA-1737.1	Contig130 (69290-69955 p)	No Hits found	
SeqID 825	SA-1738.1	Contig130 (68962-69303 p)	54	Identities = 44/104 (42%), Positives = 63/104 (60%), Gaps = 2/104 (1%) pir F81982 hypothetical protein NMA0629 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB83918.1 (AL162753) hypothetical protein NMA0629 [Neisseria meningitidis] Length = 108
SeqID 826	SA-1739.1	Contig130 (68096-68965 p)	41	Identities = 71/224 (31%), Positives = 123/224 (54%) pir G81982 hypothetical protein NMA0630 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB83919.1 (AL162753) hypothetical protein NMA0630 [Neisseria meningitidis] Length = 304
SeqID 827	SA-174.1	Contig136 (77700-78344 p)	58	Identities = 79/204 (38%), Positives = 132/204 (63%), Gaps = 4/204 (1%) sp P54176 HLY3_BACCE HEMOLYSIN III (HLY-III) pir S59967 hemolysin III - Bacillus cereus emb CAA58877.1 (X84058) novel hemolytic factor [Bacillus cereus] Length = 219

SeqID 828	SA-1740.2	Contig130 (63931-68094 p)	21	Identities = 170/765 (22%), Positives = 303/765 (39%), Gaps = 90/765 (11%) pir F81299 hypothetical protein Cj1523c [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73943.1 (AL139078) hypothetical protein Cj1523c [Campylobacter jejuni] Length = 984
SeqID 829	SA-1741.2	Contig133 (58296-58787 p)	32	Identities = 35/109 (32%), Positives = 62/109 (56%), Gaps = 1/109 (0%) pir T35305 probable acetyltransferase - Streptomyces coelicolor emb CAB51273.1 (AL096872) putative acetyltransferase [Streptomyces coelicolor A3(2)] Length = 169
SeqID 830	SA-1742.1	Contig133 (58780-60048 p)	72	Identities = 248/417 (59%), Positives = 315/417 (75%), Gaps = 4/417 (0%) pir D69981 conserved hypothetical protein yvN - Bacillus subtilis emb CAB14695.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] emb CAB14712.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 421
SeqID 831	SA-1744.1	Contig133 (60591-60896 m)	No Hits found	
SeqID 832	SA-1745.1	Contig133 (60880-61281 m)	No Hits found	
SeqID 833	SA-1746.1	Contig133 (61269-62477 m)	No Hits found	
SeqID 834	SA-1747.1	Contig133 (62479-62811 m)	No Hits found	
SeqID 835	SA-1748.1	Contig133 (62938-63327 m)	30	Identities = 26/79 (32%), Positives = 47/79 (58%), Gaps = 1/79 (1%) pir T13289 probable integrase - Streptococcus phage phi- O1205 gb AAC79517.1 (U88974) ORF1 [Streptococcus thermophilus temperate bacteriophage O1205] Length = 359
SeqID 836	SA-1749.1	Contig133 (63625-63741 m)	No Hits found	
SeqID 837	SA-175.1	Contig136 (77246-77695 p)	No Hits found	
SeqID 838	SA-1750.1	Contig133 (63820-64068 m)	No Hits found	

SeqID 839	SA-1751.1	Contig133 (64228-64722 p)	27	Identities = 28/80 (35%), Positives = 49/80 (61%) pir D69898 transcription regulator phage-related homolog yobD - Bacillus subtilis gb AAB84427.1 (AF027868) transcription regulator [Bacillus subtilis] emb CAB13777.1 (Z99114) similar to transcriptional regulator (phage-related) (Xre family) [Bacillus subtilis] Length = 112
SeqID 840	SA-1753.2	Contig103 (9236-10060 p)	64	Identities = 138/268 (51%), Positives = 186/268 (68%), Gaps = 2/268 (0%) gb AAK04909.1 AE006314_7 (AE006314) conserved hypothetical protein [Lactococcus lactis subsp. lactis] Length = 269
SeqID 841	SA-1754.1	Contig103 (8439-9236 p)	54	Identities = 91/264 (34%), Positives = 146/264 (54%), Gaps = 4/264 (1%) sp P75809 YBJI_ECOLI PROTEIN YBJI Length = 271
SeqID 842	SA-1755.1	Contig103 (4809-8348 p)	61	Identities = 458/1193 (38%), Positives = 728/1193 (60%), Gaps = 27/1193 (2%) sp P51834 SMC_BACSU CHROMOSOME PARTITION PROTEIN SMC pir G69708 chromosome segregation SMC protein - Bacillus subtilis emb CAB13467.1 (Z99112) chromosome segregation SMC protein homolog [Bacillus subtilis] Length = 1186
SeqID 843	SA-1756.1	Contig103 (4115-4801 p)	63	Identities = 115/231 (49%), Positives = 155/231 (66%), Gaps = 1/231 (0%) sp P51833 RNC_BACSU RIBONUCLEASE III (RNASE III) pir B69693 ribonuclease III (EC 3.1.26.3) - Bacillus subtilis emb CAB13466.1 (Z99112) ribonuclease III [Bacillus subtilis] Length = 249
SeqID 844	SA-1757.1	Contig103 (3571-3939 p)	54	Identities = 50/114 (43%), Positives = 72/114 (62%) emb CAC12789.1 (AJ279090) hypothetical protein [Staphylococcus carnosus] Length = 129
SeqID 845	SA-1758.1	Contig103 (2759-3568 p)	90	Identities = 223/269 (82%), Positives = 246/269 (90%) emb CAB65453.1 (AJ012051) VicX protein [Streptococcus pyogenes] Length = 270
SeqID 846	SA-1759.1	Contig103 (1406-2755 p)	86	Identities = 335/443 (75%), Positives = 392/443 (87%) emb CAB65452.1 (AJ012051) VicK protein [Streptococcus pyogenes] Length = 452

SeqID 847	SA-176.1	Contig136 (75871-77154 p)	77	Identities = 266/421 (63%), Positives = 344/421 (81%), Gaps = 3/421 (0%) gb AAG02454.1 AF290098_2 (AF290098) HMG-CoA reductase [Streptococcus pneumoniae] Length = 424
SeqID 848	SA-1760.1	Contig103 (703-1413 p)	93	Identities = 205/237 (86%), Positives = 222/237 (93%) emb CAB65451.1 (AJ012051) VicR protein [Streptococcus pyogenes] Length = 239
SeqID 849	SA-1761.1	Contig103 (15-215 p)	No Hits found	
SeqID 850	SA-1762.2	Contig118 (25352-25687 p)	57	Identities = 33/110 (30%), Positives = 65/110 (59%) gb AAC62417.1 (AF084104) hypothetical protein [Bacillus firmus] Length = 118
SeqID 851	SA-1763.1	Contig118 (24102-25268 p)	73	Identities = 227/390 (58%), Positives = 284/390 (72%), Gaps = 2/390 (0%) dbj BAB05375.1 (AP001512) chorismate synthase [Bacillus halodurans] Length = 390
SeqID 852	SA-1764.1	Contig118 (23034-24101 p)	55	Identities = 137/351 (39%), Positives = 199/351 (56%), Gaps = 4/351 (1%) sp P73997 AROB_SYNY3 3-DEHYDROQUINATE SYNTHASE pir S75507 3-dehydroquinate synthase - Synechocystis sp. (strain PCC 6803) dbj BAA18068.1 (D90911) 3-dehydroquinate synthase [Synechocystis sp.] Length = 361
SeqID 853	SA-1765.1	Contig118 (22263-22940 p)	53	Identities = 70/233 (30%), Positives = 127/233 (54%), Gaps = 12/233 (5%) sp P35146 AROD_BACSU 3-DEHYDROQUINATE DEHYDRATASE (3-DEHYDROQUINASE) (TYPE I DHQASE) pir S45563 3-dehydroquinate dehydratase (EC 4.2.1.10) aroC - Bacillus subtilis gb AAAG7501.1 (L09228) dehydroquinate dehydratase [Bacillus subtilis] emb CAB14240.1 (Z99116) 3-dehydroquinate dehydratase [Bacillus subtilis] Length = 255

SeqID 854	SA-1766.1	Contig118 (21106-22263 p)	59	Identities = 161/396 (40%), Positives = 234/396 (58%), Gaps = 11/396 (2%) sp P39587 YWBD_BACSU HYPOTHETICAL 44.4 KD PROTEIN IN EPR-GALK INTERGENIC REGION pir S39674 ywbD protein - Bacillus subtilis emb CAA51575.1 (X73124) ipa-19d [Bacillus subtilis] emb CAB15862.1 (Z99123) alternate gene name: ipa-19d~similar to hypothetical proteins [Bacillus subtilis] Length = 396
SeqID 855	SA-1767.1	Contig118 (18827-20971 m)	57	Identities = 298/682 (43%), Positives = 417/682 (60%), Gaps = 33/682 (4%) pir D69815 conserved hypothetical protein yfnl - Bacillus subtilis dbj BAA20118.1 (D86418) Yfnl [Bacillus subtilis] emb CAB12545.1 (Z99107) alternate gene name: yetP~similar to hypothetical proteins [Bacillus subtilis] Length = 653
SeqID 856	SA-1768.2	Contig118 (18227-18586 p)	91	Identities = 101/119 (84%), Positives = 110/119 (91%) gb AAK05940.1 AE006414_6 (AE006414) 50S ribosomal protein L20 [Lactococcus lactis subsp. lactis] Length = 119
SeqID 857	SA-1769.2	Contig118 (15777-15974 m)	44	Identities = 36/65 (55%), Positives = 49/65 (75%), Gaps = 2/65 (3%) gb AAK05803.1 AE006401_4 (AE006401) ferredoxin [Lactococcus lactis subsp. lactis] Length = 67
SeqID 858	SA-177.1	Contig136 (74670-75869 p)	72	Identities = 261/385 (67%), Positives = 325/385 (83%) gb AAG02448.1 AF290096_1 (AF290096) HMG-CoA synthase [Streptococcus pyogenes] Length = 391
SeqID 859	SA-1770.1	Contig118 (15299-15790 p)	40	Identities = 44/135 (32%), Positives = 66/135 (48%), Gaps = 10/135 (7%) sp P36920 EBSA_ENTFA PORE FORMING PROTEIN EBSA pir A49939 ebsA protein - Enterococcus faecalis gb AAC36851.1 (L23802) pore-forming peptide [Enterococcus faecalis] Length = 153

SeqID 860	SA-1771.1	Contig118 (14050-15270 p)	81	Identities = 274/406 (67%), Positives = 334/406 (81%), Gaps = 4/406 (0%) sp P42020 PEPT_LALC PEPTIDASE T (AMINOTRIPEPTIDASE) (TRIPETIDASE) gb AAA20627.1 (L27596) tripeptidase [Lactococcus lactis] Length = 413
SeqID 861	SA-1772.1	Contig118 (12280-13914 p)	54	Identities = 178/545 (32%), Positives = 302/545 (54%), Gaps = 26/545 (4%) pir G6992 spore cortex protein homolog ytgP - Bacillus subtilis gb AAC00276.1 (AF008220) YtgP [Bacillus subtilis] emb CAB14983.1 (Z99119) similar to spore cortex protein [Bacillus subtilis] Length = 544
SeqID 862	SA-1773.1	Contig118 (10702-12096 m)	47	Identities = 153/468 (32%), Positives = 237/468 (49%), Gaps = 23/468 (4%) dbj BAB06290.1 (AP001515) UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate ligase [Bacillus halodurans] Length = 486
SeqID 863	SA-1774.2	Contig118 (9753-10547 m)	68	Identities = 141/259 (54%), Positives = 193/259 (74%) gb AAD33527.2 AF132117_4 (AF132117) FhuA [Staphylococcus aureus] gb AAF98153.1 AF251216_1 (AF251216) FhuC [Staphylococcus aureus] Length = 265
SeqID 864	SA-1775.2	Contig118 (8797-9729 m)	55	Identities = 95/301 (31%), Positives = 178/301 (58%), Gaps = 11/301 (3%) dbj BAB07609.1 (AP001520) ferrichrome ABC transporter (ferrichrome-binding protein) [Bacillus halodurans] Length = 308
SeqID 865	SA-1776.1	Contig118 (7756-8781 m)	56	Identities = 117/313 (37%), Positives = 195/313 (61%), Gaps = 3/313 (0%) gb AAD33526.3 AF132117_3 (AF132117) ferrichrome transport permease [Staphylococcus aureus] gb AAF98154.1 AF251216_2 (AF251216) FhuB [Staphylococcus aureus] Length = 341
SeqID 866	SA-1778.1	Contig118 (6758-7759 m)	61	Identities = 122/334 (36%), Positives = 210/334 (62%), Gaps = 3/334 (0%) gb AAD33524.3 AF132117_1 (AF132117) ferrichrome transport permease [Staphylococcus aureus] gb AAF98155.1 AF251216_3 (AF251216) FhuG [Staphylococcus aureus] Length = 338

SeqID 867	SA-1779.2	Contig118 (6075-6728 p)	50	Identities = 89/233 (38%), Positives = 128/233 (54%), Gaps = 23/233 (9%) gb AAK05909.1 AE006411_4 (AE006411) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 265
SeqID 868	SA-178.1	Contig136 (73704-74543 m)	83	Identities = 215/280 (76%), Positives = 244/280 (86%), Gaps = 2/280 (0%) sp P19368 TYSY_LACLA THYMIDYLATE SYNTHASE (TS) (TSASE) pir A43797 thymidylate synthase (EC 2.1.1.45) - Lactococcus lactis subsp. lactis gb AAA25221.1 (M33770) thymidylate synthase (EC 2.1.1.45) [Lactococcus lactis] Length = 279
SeqID 869	SA-1780.2	Contig138 (28940-29971 m)	53	Identities = 137/350 (39%), Positives = 211/350 (60%), Gaps = 3/350 (0%) emb CAB54584.1 (AJ006400) histidine kinase [Streptococcus pneumoniae] Length = 365
SeqID 870	SA-1781.1	Contig138 (30022-30762 m)	51	Identities = 73/239 (30%), Positives = 127/239 (52%), Gaps = 4/239 (1%) emb CAB69806.1 (AJ243712) YVFS protein [Bacillus cereus] Length = 239
SeqID 871	SA-1783.1	Contig138 (30749-31645 m)	42	Identities = 83/216 (38%), Positives = 136/216 (62%), Gaps = 1/216 (0%) gb AAG21390.1 AF302051_4 (AF302051) ABC transporter ATP binding subunit [Bacillus licheniformis] Length = 299
SeqID 872	SA-1784.1	Contig138 (31774-33384 m)	74	Identities = 303/533 (56%), Positives = 401/533 (74%), Gaps = 15/533 (2%) gb AAF34762.1 AF228345_1 (AF228345) unknown [Listeria monocytogenes] Length = 520
SeqID 873	SA-1785.1	Contig138 (33573-34055 p)	89	Identities = 129/160 (80%), Positives = 149/160 (92%) gb AAG28749.1 AF295118_1 (AF295118) LuxS [Streptococcus pyogenes] Length = 170
SeqID 874	SA-1787.1	Contig138 (34280-35743 m)	30	Identities = 71/398 (17%), Positives = 152/398 (37%), Gaps = 16/398 (4%) gb AAD16120.1 (AF094508) dentin phosphoryn [Homo sapiens] Length = 788

SeqID 875	SA-1788.1	Contig138 (35756-36910 m)	70	Identities = 204/383 (53%), Positives = 276/383 (71%), Gaps = 3/383 (0%) sp P50840 YPSC_BACSU HYPOTHETICAL 43.5 KD PROTEIN IN COTD-KDUD INTERGENIC REGION PRECURSOR pir F69941 conserved hypothetical protein ypsC - Bacillus subtilis gb AAB38473.1 (L47838) putative [Bacillus subtilis] emb CAB14134.1 (Z99115) similar to hypothetical proteins [Bacillus subtilis] Length = 385
SeqID 876	SA-179.1	Contig136 (73130-73624 m)	62	Identities = 83/166 (50%), Positives = 121/166 (72%), Gaps = 1/166 (0%) gb AAC33872.1 (AF055727) dihydrofolate reductase [Streptococcus pneumoniae] Length = 168
SeqID 877	SA-1791.2	Contig138 (37380-37724 m)	57	Identities = 46/102 (45%), Positives = 69/102 (67%), Gaps = 14/102 (13%) sp P50839 YPSB_BACSU HYPOTHETICAL 11.6 KD PROTEIN IN COTD-KDUD INTERGENIC REGION pir E69941 hypothetical protein ypsB - Bacillus subtilis gb AAB38472.1 (L47838) putative [Bacillus subtilis] emb CAB14135.1 (Z99115) ypsB [Bacillus subtilis] Length = 98
SeqID 878	SA-1792.1	Contig86 (6483-7529 m)	57	Identities = 131/350 (37%), Positives = 200/350 (56%), Gaps = 14/350 (4%) sp P42977 PAPS_BACSU POLY(A) POLYMERASE (PAP) pir B69672 poly(A) polymerase papS - Bacillus subtilis gb AAB38446.1 (L47709) poly(A) polymerase [Bacillus subtilis] emb CAB14161.1 (Z99115) poly(A) polymerase [Bacillus subtilis] Length = 397
SeqID 879	SA-1793.1	Contig86 (4604-6472 m)	63	Identities = 269/625 (43%), Positives = 397/625 (63%), Gaps = 11/625 (1%) pir A69814 ABC transporter (ATP-binding protein) homolog yfmR - Bacillus subtilis dbj BAA20107.1 (D86418) YfmR [Bacillus subtilis] emb CAB12556.1 (Z99107) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] emb CAB12566.1 (Z99108) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 629

SeqID 880	SA-1794.1	Contig86 (4134-4607 m)	57	Identities = 56/161 (34%), Positives = 92/161 (56%), Gaps = 8/161 (4%) sp P39337 YJGM_ECOLI_HYPOTHETICAL 18.6 KDA PROTEIN IN ARG1-VALS INTERGENIC REGION Length = 167
SeqID 881	SA-1795.1	Contig86 (2328-4067 m)	58	Identities = 227/565 (40%), Positives = 342/565 (60%), Gaps = 1/565 (0%) emb CAB69751.1 (AL137187) putative ABC transporter [Streptomyces coelicolor A3(2)] Length = 577
SeqID 882	SA-1796.1	Contig86 (554-2245 m)	65	Identities = 270/611 (44%), Positives = 391/611 (63%), Gaps = 31/611 (5%) emb CAB69752.1 (AL137187) putative ABC transporter [Streptomyces coelicolor A3(2)] Length = 642
SeqID 883	SA-1797.1	Contig86 (4-513 p)	69	Identities = 78/167 (46%), Positives = 118/167 (69%), Gaps = 1/167 (0%) pir A33141 hypothetical protein (gtfD 3 region) - Streptococcus mutans Length = 169
SeqID 884	SA-1798.3	Contig93 (7851-9347 p)	65	Identities = 250/500 (50%), Positives = 329/500 (65%), Gaps = 4/500 (0%) sp P29851 MALQ_STRPN 4-ALPHA-GLUCANOTRANSFERASE (AMYLOMALTASE) (DISPROPORTIONATING ENZYME) (D-ENZYME) gb AAA26923.1 (J01796) amylomaltase [Streptococcus pneumoniae] Length = 505
SeqID 885	SA-1799.3	Contig93 (9359-11623 p)	59	Identities = 297/776 (38%), Positives = 452/776 (57%), Gaps = 41/776 (5%) sp P39123 PHSG_BACSU GLYCOGEN PHOSPHORYLASE pir S36628 glycogen phosphorylase (EC 2.4.1.1) - Bacillus subtilis pir S40052 glycogen phosphorylase (EC 2.4.1.1) glgP - Bacillus subtilis emb CAA81044.1 (Z25795) Glycogen Phosphorylase [Bacillus subtilis] gb AAC00218.1 (AF008220) glycogen phosphorylase [Bacillus subtilis] emb CAB15072.1 (Z99119) glycogen phosphorylase [Bacillus subtilis] Length = 798

SeqID 886	SA-18.1	Contig137 (25957-28218 p)	35	Identities = 178/715 (24%), Positives = 273/715 (37%), Gaps = 117/715 (16%) pir G41662 130K surface exclusion protein Sec10 precursor - Enterococcus faecalis plasmid pCF10 gb AAA65847.1 (M64978) surface exclusion protein [Plasmid pCF10] Length = 891
SeqID 887	SA-1801.2	Contig139 (45477-45692 p)	75	Identities = 28/76 (36%), Positives = 55/76 (71%), Gaps = 4/76 (5%) pir E83139 exodeoxyribonuclease VII small subunit PA4042 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG07429.1 AE004821_2 (AE004821) exodeoxyribonuclease VII small subunit [Pseudomonas aeruginosa] Length = 80
SeqID 888	SA-1802.1	Contig139 (44159-45499 p)	63	Identities = 194/446 (43%), Positives = 284/446 (63%), Gaps = 10/446 (2%) sp P54521 EX7L_BACSU PUTATIVE EXODEOXYRIBONUCLEASE LARGE SUBUNIT (EXONUCLEASE VII LARGE SUBUNIT) pir G69960 exodeoxyribonuclease VII (large subunit) homolog yqiB - Bacillus subtilis dbj BAA12573.1 (D84432) YqiB [Bacillus subtilis] emb CAB14361.1 (Z99116) similar to exodeoxyribonuclease VII (large subunit) [Bacillus subtilis] Length = 448
SeqID 889	SA-1803.1	Contig139 (43197-44033 p)	67	Identities = 149/277 (53%), Positives = 191/277 (68%) sp P96051 YFOL_STRTR HYPOTHETICAL 29.9 KD PROTEIN IN FOLD-PBP2B INTERGENIC REGION (ORF1091) gb AAC44613.1 (U58210) orf1091 [Streptococcus thermophilus] Length = 278
SeqID 890	SA-1804.1	Contig139 (42346-43200 p)	83	Identities = 209/282 (74%), Positives = 248/282 (87%) sp P96050 FOLD_STRTR FOLD BIFUNCTIONAL PROTEIN [INCLUDES: METHYLENETETRAHYDROFOLATE DEHYDROGENASE; METHENYL-TETRAHYDROFOLATE CYCLOHYDROLASE] gb AAC44612.1 (U58210) tetrahydrofolate dehydrogenase/cyclohydrolase [Streptococcus thermophilus] Length = 284

SeqID 891	SA-1806.1	Contig139 (40513-42207 p)	60	Identities = 238/548 (43%), Positives = 343/548 (62%), Gaps = 14/548 (2%) dbj BAB04825.1 (AP001510) phosphomannomutase [Bacillus halodurans] Length = 578
SeqID 892	SA-1807.2	Contig139 (39492-40226 p)	78	Identities = 153/239 (64%), Positives = 193/239 (80%) ref NP_069514.1 glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus] pir H69334 glutamine transport protein glnQ - Archaeoglobus fulgidus gb AAB90561.1 (AE001058) glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus] Length = 242
SeqID 893	SA-1808.2	Contig139 (38801-39499 p)	55	Identities = 92/209 (44%), Positives = 129/209 (61%), Gaps = 10/209 (4%) ref NP_069070.1 glutamine ABC transporter, permease protein (glnP) [Archaeoglobus fulgidus] pir H69278 glutamine ABC transporter, permease protein (glnP) homolog - Archaeoglobus fulgidus gb AAB91000.1 (AE001090) glutamine ABC transporter, permease protein (glnP) [Archaeoglobus fulgidus] Length = 224
SeqID 894	SA-1809.2	Contig89 (8855-9556 m)	68	Identities = 150/231 (64%), Positives = 178/231 (76%), Gaps = 1/231 (0%) sp Q9ZHA7 DCOP_STRPN OROTIDINE 5 - PHOSPHATE DECARBOXYLASE (OMP DECARBOXYLASE) (OMPDCase) gb AAC95452.1 (AF068902) orotidine-5 - decarboxylase PyrF [Streptococcus pneumoniae] Length = 233
SeqID 895	SA-181.1	Contig136 (71686-72912 m)	85	Identities = 306/395 (77%), Positives = 357/395 (89%), Gaps = 1/395 (0%) gb AAF63738.1 AF236863_2 (AF236863) protease ClpX [Lactococcus lactis] Length = 411
SeqID 896	SA-1810.1	Contig89 (8213-8842 m)	84	Identities = 152/208 (73%), Positives = 180/208 (86%) gb AAC95453.1 (AF068902) orotate phosphoribosyltransferase PyrE [Streptococcus pneumoniae] Length = 210

SeqID 897	SA-1812.1	Contig89 (6909-8201 m)	57	Identities = 176/422 (41%), Positives = 255/422 (59%), Gaps = 8/422 (1%) sp O66990 PYRC_AQUAE DIHYDROOROTASE (DHOASE) pir C70370 dihydroorotase - Aquifex aeolicus gb AAC06948.1 (AE000708) dihydroorotase [Aquifex aeolicus] Length = 422
SeqID 898	SA-1813.1	Contig89 (5821-6744 m)	80	Identities = 197/303 (65%), Positives = 250/303 (82%) gb AAF72727.1 (AF264709) aspartate transcarbamoylase [Enterococcus faecalis] Length = 308
SeqID 899	SA-1814.1	Contig89 (4731-5807 m)	82	Identities = 242/355 (68%), Positives = 305/355 (85%) emb CAB89872.1 (AJ132624) carbamoyl phosphate synthetase small subunit [Lactococcus lactis] Length = 357
SeqID 900	SA-1815.1	Contig89 (1518-4700 m)	84	Identities = 771/1062 (72%), Positives = 900/1062 (84%), Gaps = 5/1062 (0%) emb CAA03928.1 (AJ000109) carbamoylphosphate synthetase [Lactococcus lactis] Length = 1064
SeqID 901	SA-1816.1	Contig89 (1068-1391 m)	No Hits found	
SeqID 902	SA-1817.2	Contig89 (289-1071 m)	22	Identities = 39/130 (30%), Positives = 59/130 (45%), Gaps = 26/130 (20%) pir T38271 conserved hypothetical protein SPAC23C4.16c - fission yeast (Schizosaccharomyces pombe) emb CAB16887.1 (Z99753) conserved hypothetical protein [Schizosaccharomyces pombe] Length = 424
SeqID 903	SA-1818.2	Contig91 (6891-7685 p)	68	Identities = 166/266 (62%), Positives = 207/266 (77%), Gaps = 2/266 (0%) gb AAK05380.1 AE006360_9 (AE006360) glutamate racemase [EC 5.1.1.3] [Lactococcus lactis subsp. lactis] Length = 271
SeqID 904	SA-1819.1	Contig91 (7682-8656 p)	38	Identities = 96/189 (50%), Positives = 130/189 (67%), Gaps = 1/189 (0%) pir C69986 conserved hypothetical protein ysnA - Bacillus subtilis emb CAA99555.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14796.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 198

SeqID 905	SA-182.1	Contig136 (71079-71675 m)	80	Identities = 141/193 (73%), Positives = 165/193 (85%) gb AAF63739.1 AF236863_3 (AF236863) hypothetical GTP-binding protein [Lactococcus lactis] Length = 195
SeqID 906	SA-1820.1	Contig91 (8638-9159 p)	53	Identities = 55/169 (32%), Positives = 96/169 (56%), Gaps = 1/169 (0%) dbj BAB06785.1 (AP001517) unknown conserved protein [Bacillus halodurans] Length = 169
SeqID 907	SA-1821.1	Contig91 (9156-9629 p)	44	Identities = 37/136 (27%), Positives = 73/136 (53%), Gaps = 1/136 (0%) dbj BAB05201.1 (AP001512) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 146
SeqID 908	SA-1822.1	Contig91 (9748-10359 p)	43	Identities = 66/271 (24%), Positives = 116/271 (42%), Gaps = 35/271 (12%) dbj BAB05248.1 (AP001512) integrase/recombinase [Bacillus halodurans] Length = 299
SeqID 909	SA-1823.1	Contig91 (10398-10685 p)	70	Identities = 52/106 (49%), Positives = 78/106 (73%) pir A83244 conserved hypothetical protein PA3198 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG06586.1 AE004744_1 (AE004744) conserved hypothetical protein [Pseudomonas aeruginosa] Length = 250
SeqID 910	SA-1824.1	Contig91 (10786-11070 p)	49	Identities = 39/138 (28%), Positives = 65/138 (46%), Gaps = 14/138 (10%) sp P35154 YPUG_BACSU HYPOTHETICAL 29.6 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX7) pir S45549 conserved hypothetical protein ypuG - Bacillus subtilis gb AA67487.1 (L09228) ORFX7 [Bacillus subtilis] emb CAB14254.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] Length = 251
SeqID 911	SA-1825.1	Contig91 (11067-11651 p)	52	Identities = 75/180 (41%), Positives = 107/180 (58%), Gaps = 10/180 (5%) dbj BAB05280.1 (AP001512) unknown conserved protein [Bacillus halodurans] Length = 197
SeqID 912	SA-1826.1	Contig91 (11641-12363 p)	67	Identities = 130/239 (54%), Positives = 175/239 (72%), Gaps = 2/239 (0%) dbj BAB05295.1 (AP001512) pseudouridylylate synthase [Bacillus halodurans] Length = 242

SeqID 913	SA-1827.1	Contig91 (12363-12614 p)	66	Identities = 43/75 (57%), Positives = 57/75 (75%), Gaps = 1/75 (1%) pir G72251 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD36530.1 AE001797_10 (AE001797) conserved hypothetical protein [Thermotoga maritima] Length = 81
SeqID 914	SA-1828.1	Contig91 (12762-12881 p)	No Hits found	
SeqID 915	SA-183.1	Contig136 (70469-71098 p)	36	Identities = 44/172 (25%), Positives = 80/172 (45%), Gaps = 27/172 (15%) gb AAG09969.1 AF248037_4 (AF248037) unknown [Streptococcus agalactiae] Length = 174
SeqID 916	SA-1830.1	Contig91 (12638-14077 m)	53	Identities = 165/480 (34%), Positives = 261/480 (54%), Gaps = 10/480 (2%) ref NP_069673.1 TRK potassium uptake system protein (trkH) [Archaeoglobus fulgidus] pir G69354 TRK potassium uptake system protein (trkH) homolog - Archaeoglobus fulgidus gb AAB90400.1 (AE001046) TRK potassium uptake system protein (trkH) [Archaeoglobus fulgidus] Length = 478
SeqID 917	SA-1832.1	Contig91 (14082-15431 m)	52	Identities = 136/446 (30%), Positives = 238/446 (52%), Gaps = 12/446 (2%) ref NP_069672.1 TRK potassium uptake system protein (trkA-2) [Archaeoglobus fulgidus] sp O29420 TRKA_ARCFU TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA HOMOLOG pir F69354 TRK potassium uptake system protein (trkA-2) homolog - Archaeoglobus fulgidus gb AAB90401.1 (AE001046) TRK potassium uptake system protein (trkA-2) [Archaeoglobus fulgidus] Length = 436
SeqID 918	SA-1834.1	Contig85 (10632-11486 p)	50	Identities = 91/289 (31%), Positives = 146/289 (50%), Gaps = 14/289 (4%) sp P12055 STR_STAAU STREPTOMYCIN RESISTANCE PROTEIN pir S00938 str protein - Staphylococcus aureus plasmid pS194 emb CAA29839.1 (X06627) ORF (str) [Staphylococcus aureus] Length = 282
SeqID 919	SA-1835.1	Contig85 (9974-10453 p)	No Hits found	
SeqID 920	SA-1836.1	Contig85 (8970-9626 p)	No Hits found	

SeqID 921	SA-1837.1	Contig85 (8753-9031 p)	45	Identities = 31/103 (30%), Positives = 57/103 (55%), Gaps = 4/103 (3%) pir G75166 hypothetical protein PAB0331 - Pyrococcus abyssi (strain Orsay) emb CAB49414.1 (AJ248284) hypothetical protein [Pyrococcus abyssi] Length = 114
SeqID 922	SA-1838.1	Contig85 (7686-8174 p)	26	Identities = 32/110 (29%), Positives = 54/110 (49%), Gaps = 4/110 (3%) ref NP_057431.1 putative N-acetyltransferase Camello 2 [Homo sapiens] gb AAAF22299.1 AF185571_1 (AF185571) putative N-acetyltransferase Camello 2 [Homo sapiens] Length = 227
SeqID 923	SA-1839.1	Contig85 (6726-7517 p)	35	Identities = 42/188 (22%), Positives = 94/188 (49%), Gaps = 4/188 (2%) pir B70082 hypothetical protein yxIG - Bacillus subtilis dbj BAA11738.1 (D83026) hypothetical [Bacillus subtilis] emb CAB15891.1 (Z99123) yxIG [Bacillus subtilis] Length = 259
SeqID 924	SA-1840.1	Contig85 (5857-6759 p)	58	Identities = 105/301 (34%), Positives = 176/301 (57%), Gaps = 11/301 (3%) sp P54592 YHCH_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLPD-CSPB INTERGENIC REGION pir D69822 ABC transporter (ATP-binding protein) homolog yhcH - Bacillus subtilis emb CAA65691.1 (X96983) hypothetical protein [Bacillus subtilis] emb CAB12736.1 (Z99108) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 305
SeqID 925	SA-1841.1	Contig85 (4576-5652 p)	19	Identities = 39/134 (29%), Positives = 71/134 (52%), Gaps = 16/134 (11%) gb AAB71491.1 (U53767) ORF6 [Bacillus pumilus] Length = 211
SeqID 926	SA-1842.1	Contig85 (3704-4327 p)	50	Identities = 50/186 (26%), Positives = 105/186 (55%), Gaps = 5/186 (2%) pir S32217 hypothetical protein 2 - Bacillus megaterium emb CAA79986.1 (Z21972) ORF2 [Bacillus megaterium] Length = 216

SeqID 927	SA-1843.1	Contig85 (3203-3499 m)	58	Identities = 53/96 (55%), Positives = 70/96 (72%) pir C70033 hypothetical protein yvdC - Bacillus subtilis emb CAB15470.1 (Z99121) yvdC [Bacillus subtilis] Length = 106
SeqID 928	SA-1844.1	Contig85 (2626-3186 m)	62	Identities = 83/186 (44%), Positives = 117/186 (62%) dbj BAB06803.1 (AP001517) unknown conserved protein [Bacillus halodurans] Length = 187
SeqID 929	SA-1845.1	Contig85 (1458-2333 m)	21	Identities = 40/132 (30%), Positives = 63/132 (47%), Gaps = 8/132 (6%) pir E69787 hypothetical protein ydlL - Bacillus subtilis dbj BAA19725.1 (D88802) transmembrane [Bacillus subtilis] emb CAB12420.1 (Z99107) ydlL [Bacillus subtilis] Length = 244
SeqID 930	SA-1846.1	Contig85 (999-1301 p)	47	Identities = 38/92 (41%), Positives = 53/92 (57%), Gaps = 1/92 (1%) sp P54510 YQHL_BACSU HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIA INTERGENIC REGION pir C69959 glpE protein homolog yqhL - Bacillus subtilis dbj BAA12549.1 (D84432) YqhL [Bacillus subtilis] emb CAB14385.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] Length = 126
SeqID 931	SA-1847.1	Contig85 (59-826 p)	96	Identities = 252/255 (98%), Positives = 252/255 (98%) emb CAA51283.1 (X72754) cAMP factor [Streptococcus agalactiae] Length = 255
SeqID 932	SA-1849.1	Contig88 (13289-14137 p)	54	Identities = 94/276 (34%), Positives = 154/276 (55%), Gaps = 2/276 (0%) dbj BAB07346.1 (AP001519) unknown conserved protein [Bacillus halodurans] Length = 283
SeqID 933	SA-185.1	Contig136 (69874-70248 m)	No Hits found	
SeqID 934	SA-1850.1	Contig88 (12909-13289 p)	No Hits found	
SeqID 935	SA-1851.1	Contig88 (12304-12837 p)	No Hits found	
SeqID 936	SA-1852.1	Contig88 (11759-12253 p)	61	Identities = 65/165 (39%), Positives = 102/165 (61%), Gaps = 5/165 (3%) gb AAAF66143.1 (U81488) TcsEorf2 [Lactococcus lactis subsp. cremoris] Length = 171

SeqID 937	SA-1853.1	Contig88 (10706-11140 m)	37	Identities = 42/134 (31%), Positives = 55/134 (40%), Gaps = 18/134 (13%) dbj BAA99921.1 (AP001306) contains similarity to cell wall-plasma membrane linker protein~gene_id:MKA23.5 [Arabidopsis thaliana] Length = 1480
SeqID 938	SA-1855.1	Contig88 (9606-11570 p)	66	Identities = 320/660 (48%), Positives = 439/660 (66%), Gaps = 46/660 (6%) dbj BAB04547.1 (AP001510) PTS system, fructose-specific enzyme II, BC component [Bacillus halodurans] Length = 625
SeqID 939	SA-1858.1	Contig88 (8698-9609 p)	61	Identities = 146/303 (48%), Positives = 197/303 (64%) sp O31714 K1PF_BACSU 1-PHOSPHOFRUCTOKINASE (FRUCTOSE 1-PHOSPHATE KINASE) pir JA69627 fructose 1-phosphate kinase fruB - Bacillus subtilis emb CAB13312.1 (Z99111) fructose-1-phosphate kinase [Bacillus subtilis] gb AAC24914.1 (AF012285) fructose-1-phosphate kinase [Bacillus subtilis] Length = 303
SeqID 940	SA-1859.1	Contig88 (7958-8701 p)	53	Identities = 96/248 (38%), Positives = 148/248 (58%), Gaps = 1/248 (0%) dbj BAB04545.1 (AP001510) transcriptional repressor [Bacillus halodurans] Length = 251
SeqID 941	SA-186.1	Contig136 (69066-69623 m)	No Hits found	
SeqID 942	SA-1860.1	Contig88 (6592-7827 p)	68	Identities = 216/410 (52%), Positives = 284/410 (68%) emb CAB89121.1 (AJ277485) beta-lactam resistance factor [Streptococcus pneumoniae] Length = 410
SeqID 943	SA-1861.1	Contig88 (4808-6448 p)	21	Identities = 93/183 (50%), Positives = 116/183 (62%), Gaps = 9/183 (4%) pir C60328 hypothetical protein 2 (sr 5 region) - Streptococcus mutans (strain OMZ175, serotype f) Length = 179
SeqID 944	SA-1862.1	Contig88 (3679-4602 p)	48	Identities = 85/301 (28%), Positives = 150/301 (49%), Gaps = 7/301 (2%) pir H75119 probable 2-dehydropantoate 2-reductase (EC 1.1.1.169) PAB0512 - Pyrococcus abyssi (strain Orsay) emb CAB49673.1 (AJ248285) PROBABLE 2-DEHYDROPANTOATE 2-REDUCTASE (EC 1.1.1.169) [Pyrococcus abyssi] Length = 300

SeqID 945	SA-1863.1	Contig88 (3120-3626 m)	No Hits found	
SeqID 946	SA-1866.2	Contig129 (15399-16343 m)	54	Identities = 111/300 (37%), Positives = 175/300 (58%), Gaps = 4/300 (1%) sp P37517 CCPB_BACSU CATABOLITE CONTROL PROTEIN B pir S66011 transcription regulator homolog yyaG - Bacillus subtilis dbj BAA05217.1 (D26185) LACI family of transcriptional repressor (probable) [Bacillus subtilis] emb CAB16124.1 (Z99124) similar to transcriptional regulator (LacI family) [Bacillus subtilis] Length = 311
SeqID 947	SA-1867.1	Contig129 (14405-15340 p)	31	Identities = 63/218 (28%), Positives = 107/218 (48%), Gaps = 3/218 (1%) gb AAB17013.1 (L38252) esterase [Acinetobacter lwoffii] Length = 303
SeqID 948	SA-1868.1	Contig129 (13444-14127 m)	90	Identities = 228/228 (100%), Positives = 228/228 (100%) emb CAA72897.1 (Y12224) hypothetical protein [Streptococcus agalactiae] Length = 228
SeqID 949	SA-1869.1	Contig129 (12353-13390 m)	95	Identities = 343/345 (99%), Positives = 343/345 (99%) emb CAA72898.1 (Y12224) hypothetical protein [Streptococcus agalactiae] Length = 345
SeqID 950	SA-1870.1	Contig129 (11665-12273 m)	97	Identities = 202/202 (100%), Positives = 202/202 (100%) sp O54086 SODM_STRAG SUPEROXIDE DISMUTASE [MN-FE] Length = 202
SeqID 951	SA-1871.1	Contig129 (10476-11327 m)	63	Identities = 118/275 (42%), Positives = 183/275 (65%) dbj BAA11696.1 (D83026) LicT antiterminator [Bacillus subtilis] emb CAB15944.1 (Z99124) transcriptional antiterminator (BglG family) [Bacillus subtilis] Length = 277

SeqID 952	SA-1872.2	Contig129 (8615-10483 m)	55	Identities = 225/594 (37%), Positives = 351/594 (58%), Gaps = 11/594 (1%) sp P40739 PTBA_BACSU PTS SYSTEM, BETA-GLUCOSIDES-SPECIFIC IIABC COMPONENT (EIABC-BGL) (BETA-GLUCOSIDES-PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC COMPONENT) (EII-BGL) pir 40406 beta-glucoside permease - Bacillus subtilis emb CAA84286.1 (Z34526) beta-glucoside permease [Bacillus subtilis] Length = 609
SeqID 953	SA-1874.1	Contig115 (24509-25441 p)	77	Identities = 195/308 (63%), Positives = 245/308 (79%), Gaps = 4/308 (1%) sp P24137 OPPF_BACSU OLIGOPEPTIDE TRANSPORT ATP-BINDING PROTEIN OPF pir E38447 oligopeptide transport ATP-binding protein oppF - Bacillus subtilis gb AAA62692.1 (M57689) sporulation protein [Bacillus subtilis] Length = 308
SeqID 954	SA-1875.1	Contig115 (23463-24509 p)	80	Identities = 230/342 (67%), Positives = 283/342 (82%), Gaps = 2/342 (0%) gb AAF73093.1 AF103793_4 (AF103793) ATPase OppD [Listeria monocytogenes] Length = 358
SeqID 955	SA-1876.1	Contig115 (22419-23450 p)	63	Identities = 157/325 (48%), Positives = 219/325 (67%), Gaps = 4/325 (1%) gb AAF73092.1 AF103793_3 (AF103793) transmembrane protein OppC [Listeria monocytogenes] Length = 344
SeqID 956	SA-1878.1	Contig115 (21495-22409 p)	69	Identities = 147/304 (48%), Positives = 221/304 (72%), Gaps = 1/304 (0%) gb AAF73091.1 AF103793_2 (AF103793) transmembrane protein OppB [Listeria monocytogenes] Length = 309
SeqID 957	SA-1879.2	Contig115 (19721-21376 p)	74	Identities = 352/512 (68%), Positives = 416/512 (80%), Gaps = 2/512 (0%) gb AAD17886.1 (AF100456) hyaluronate-associated protein precursor [Streptococcus equi] Length = 522

SeqID 958	SA-188.1	Contig136 (67560-68948 m)	74	Identities = 278/454 (61%), Positives = 351/454 (77%), Gaps = 3/454 (0%) pir A69751 histidine permease homolog ybgF - Bacillus subtilis emb CAB12034.1 (Z99105) similar to histidine permease [Bacillus subtilis] dbj BAA33138.1 (AB006424) ybgF [Bacillus subtilis] Length = 470
SeqID 959	SA-1881.2	Contig132 (48543-48848 p)	78	Identities = 89/101 (88%), Positives = 94/101 (92%) sp Q9WVW6 RL24_STRPN 50S RIBOSOMAL PROTEIN L24 gb AAD33267.1 AF126059_8 (AF126059) Rpl24 [Streptococcus pneumoniae] gb AAD33276.1 (AF126060) Rpl24 [Streptococcus pneumoniae] gb AAD33285.1 (AF126061) Rpl24 [Streptococcus pneumoniae] Length = 101
SeqID 960	SA-1882.2	Contig132 (48872-49414 p)	90	Identities = 157/180 (87%), Positives = 172/180 (95%) gb AAK06185.1 AE006437_15 (AE006437) 50S ribosomal protein L5 [Lactococcus lactis subsp. lactis] Length = 180
SeqID 961	SA-1883.1	Contig132 (49772-50170 p)	86	Identities = 100/132 (75%), Positives = 116/132 (87%) sp P12879 RS8_BACSU 30S RIBOSOMAL PROTEIN S8 (BS8) gb AAB06813.1 (L47971) ribosomal protein S8 [Bacillus subtilis] Length = 132
SeqID 962	SA-1884.1	Contig132 (50280-50816 p)	72	Identities = 110/178 (61%), Positives = 134/178 (74%) sp P02391 RL6_BACST 50S RIBOSOMAL PROTEIN L6 (BL10) Length = 178
SeqID 963	SA-1885.1	Contig132 (50917-51273 p)	78	Identities = 86/120 (71%), Positives = 97/120 (80%), Gaps = 2/120 (1%) sp P46899 RL18_BACSU 50S RIBOSOMAL PROTEIN L18 gb AAB06815.1 (L47971) ribosomal protein L18 [Bacillus subtilis] Length = 120
SeqID 964	SA-1886.1	Contig132 (51292-51786 p)	81	Identities = 119/158 (75%), Positives = 139/158 (87%) sp P02357 RS5_BACST 30S RIBOSOMAL PROTEIN S5 (BS5) pir R3BS5F ribosomal protein S5 - Bacillus stearothermophilus gb AAA22699.1 (M57621) ribosomal protein S5 [Bacillus stearothermophilus] Length = 166
SeqID 965	SA-1887.1	Contig132 (51534-51890 m)	No Hits found	

SeqID 966	SA-1888.1	Contig132 (51801-51980 p)	67	Identities = 40/58 (68%), Positives = 46/58 (78%) sp O06444 RL30_STAAU 50S RIBOSOMAL PROTEIN L30 gb AAB54020.1 (U96620) ribosomal protein L30 [Staphylococcus aureus] Length = 59
SeqID 967	SA-1891.1	Contig132 (52105-52545 p)	84	Identities = 116/146 (79%), Positives = 128/146 (87%) sp O06445 RL15_STAAU 50S RIBOSOMAL PROTEIN L15 gb AAB54021.1 (U96620) ribosomal protein L15 [Staphylococcus aureus] Length = 146
SeqID 968	SA-1892.1	Contig132 (52566-53870 p)	82	Identities = 292/433 (67%), Positives = 361/433 (82%), Gaps = 2/433 (0%) sp P27148 SECY_LACLA PREPROTEIN TRANSLOCASE SECY SUBUNIT pir S17985 preprotein translocase secY - Lactococcus lactis subsp. lactis emb CAA41939.1 (X59250) SecY protein [Lactococcus lactis] prf 1715214A secY gene [Lactococcus lactis] Length = 439
SeqID 969	SA-1893.1	Contig132 (53965-54603 p)	77	Identities = 146/214 (68%), Positives = 170/214 (79%), Gaps = 6/214 (2%) sp P27143 KAD_LACLA ADENYLATE KINASE (ATP- AMP TRANSPHOSPHORYLASE) pir S17987 adenylate kinase (EC 2.7.4.3) - Lactococcus lactis subsp. lactis pir B44812 adenylate kinase (EC 2.7.4.3) - Lactococcus lactis emb CAA41940.1 (X59250) adenylate kinase [Lactococcus lactis] Length = 215

SeqID 970	SA-1894.2	Contig132 (28026-29129 p)	23	<p>Identities = 51/213 (23%), Positives = 89/213 (40%), Gaps = 32/213 (15%) pir T19214 UDP-glucose--glycoprotein glucosyltransferase (EC 2.4.1.-) precursor F26H9.8 - Caenorhabditis elegans emb CAB03874.1 (Z81467) predicted using Genefinder~Similarity to Drosophila UDP-glucose:glycoprotein glucosyltransferase (TR:Q09332), contains similarity to Pfam domain: PF01501 (Glycosyl transferase family 8), Score=-22.6, E-value=4.8e-05, N=1~cDNA EST yk250b10.3 co> emb CAB04207.1 (Z81516) predicted using Genefinder~Similarity to Drosophila UDP-glucose:glycoprotein glucosyltransferase (TR:Q09332), contains similarity to Pfam domain: PF01501 (Glycosyl transferase family 8), Score=-22.6, E-value=4.8e-05, N=1~cDNA EST yk250b10.3 co> Length = 1377</p>
SeqID 971	SA-1897.1	Contig132 (29215-30513 p)	83	<p>Identities = 324/433 (74%), Positives = 367/433 (83%) dbj BAB04344.1 (AP001509) adenylosuccinate lyase [Bacillus halodurans] Length = 433</p>
SeqID 972	SA-1898.1	Contig132 (30700-31569 p)	No Hits found	
SeqID 973	SA-1899.1	Contig132 (31858-32856 p)	74	<p>Identities = 196/322 (60%), Positives = 254/322 (78%) sp O32055 RUVB_BACSU HOLLIDAY JUNCTION DNA HELICASE RUVB emb CAB75331.1 (Y15896) RuvB protein [Bacillus subtilis] Length = 334</p>
SeqID 974	SA-190.1	Contig136 (66611-67555 m)	53	<p>Identities = 140/305 (45%), Positives = 191/305 (61%), Gaps = 6/305 (1%) sp Q47690 MMUM_ECOLI HOMOCYSTEINE S-METHYLTRANSFERASE (S-METHYLMETHIONINE:HOMOCYSTEINE METHYLTRANSFERASE) pir E64751 probable membrane protein yagD - Escherichia coli gb AAB08682.1 (U70214) similar to S. cerevisiae YLL062c [Escherichia coli] gb AAC73364.1 (AE000134) putative enzyme [Escherichia coli K12] Length = 310</p>

SeqID 975	SA-1900.3	Contig132 (33008-33445 p)	61	Identities = 76/141 (53%), Positives = 100/141 (70%), Gaps = 4/141 (2%) gb AAK06300.1 AE006449_6 (AE006449) protein-tyrosine phosphatase (EC 3.1.3.48) [Lactococcus lactis subsp. lactis] Length = 145
SeqID 976	SA-1901.2	Contig106 (12011-12379 m)	62	Identities = 51/116 (43%), Positives = 79/116 (67%) dbj BAA83965.1 (AB024564) YHCF [Bacillus halodurans] dbj BAB04102.1 (AP001508) transcriptional regulator (GntR family) [Bacillus halodurans] Length = 126
SeqID 977	SA-1902.2	Contig106 (12524-15628 p)	78	Identities = 647/1036 (62%), Positives = 814/1036 (78%), Gaps = 4/1036 (0%) gb AAF98350.1 (AF280766) DNA polymerase III alpha subunit [Streptococcus pyogenes] Length = 1034
SeqID 978	SA-1904.1	Contig106 (15709-16731 p)	88	Identities = 270/337 (80%), Positives = 302/337 (89%), Gaps = 1/337 (0%) gb AAF25803.1 AF172173_1 (AF172173) phosphofructokinase [Streptococcus thermophilus] Length = 339
SeqID 979	SA-1906.1	Contig106 (16780-18282 p)	89	Identities = 413/500 (82%), Positives = 452/500 (89%) gb AAF25804.1 AF172173_2 (AF172173) pyruvate kinase [Streptococcus thermophilus] Length = 500
SeqID 980	SA-1908.2	Contig106 (18453-19010 p)	48	Identities = 62/185 (33%), Positives = 97/185 (51%), Gaps = 12/185 (6%) sp P72365 LEP_STAAU SIGNAL PEPTIDASE IB (SPASE IB) (LEADER PEPTIDASE IB) gb AAC44435.1 (U65000) type-I signal peptidase SpsB [Staphylococcus aureus] Length = 191
SeqID 981	SA-1909.2	Contig133 (13473-13898 m)	78	Identities = 92/141 (65%), Positives = 115/141 (81%) sp P23494 LACA_LACLA GALACTOSE-6-PHOSPHATE ISOMERASE LACA SUBUNIT pir A39778 galactose-6-phosphate isomerase (EC 5.3.1.-) component LacA - Lactococcus lactis gb AA25168.1 (M65190) lacA [Lactococcus lactis] gb AAA25177.1 (M60447) galactose 6-P isomerase [Lactococcus lactis] Length = 141

SeqID 982	SA-191.2	Contig136 (64409-66517 p)	72	Identities = 437/589 (74%), Positives = 513/589 (86%), Gaps = 5/589 (0%) gb AAA68910.1 (L34677) Clp-like ATP-dependent protease binding subunit [Bos taurus] Length = 586
SeqID 983	SA-1910.1	Contig133 (12937-13452 m)	86	Identities = 138/171 (80%), Positives = 157/171 (91%) sp P23495 LACB_LACLA GALACTOSE-6-PHOSPHATE ISOMERASE LACB SUBUNIT pir B39778 galactose-6-phosphate isomerase (EC 5.3.1.-) component LacB - Lactococcus lactis gb AAA25169.1 (M65190) lacB [Lactococcus lactis] gb AAA25178.1 (M60447) galactose 6-P isomerase [Lactococcus lactis] Length = 171
SeqID 984	SA-1911.1	Contig133 (11994-12926 m)	72	Identities = 192/310 (61%), Positives = 236/310 (75%) sp P23391 LACC_LACLA TAGATOSE-6-PHOSPHATE KINASE (PHOSPHOTAGATOKINASE) pir C39778 tagatose-6-phosphate kinase (EC 2.7.1.-) LacC - Lactococcus lactis gb AAA25170.1 (M65190) lacC [Lactococcus lactis] gb AAA25179.1 (M60447) tagatose 6-P kinase [Lactococcus lactis] Length = 310
SeqID 985	SA-1912.1	Contig133 (11015-11992 m)	89	Identities = 253/325 (77%), Positives = 295/325 (89%) sp P26593 LACD_LACLA TAGATOSE 1,6-DIPHOSPHATE ALDOLASE (TAGATOSE-BISPHOSPHATE ALDOLASE) (D-TAGATOSE-1,6-BISPHOSPHATE ALDOLASE) pir D39778 LacD tagatose-1,6-diphosphate aldolase - Lactococcus lactis gb AAA25171.1 (M65190) lacD [Lactococcus lactis] gb AAA25180.1 (M60447) tagatose 1,6-diP aldolase [Lactococcus lactis] Length = 326
SeqID 986	SA-1913.1	Contig133 (10065-10961 m)	66	Identities = 173/298 (58%), Positives = 219/298 (73%) sp P23496 LAXP_LACLA LACX PROTEIN, PLASMID pir D23696 lacX protein - Lactococcus lactis gb AAA25184.1 (M60447) ORF [Lactococcus lactis] Length = 299

SeqID 987	SA-1915.1	Contig133 (9119-9970 m)	47	Identities = 101/278 (36%), Positives = 160/278 (57%), Gaps = 1/278 (0%) sp Q54087 LRPR_STREQ LEUCINE RICH PROTEIN pir S39972 leucine-rich protein - Streptococcus equisimilis emb CAA51350.1 (X72832) leucine rich protein [Streptococcus equisimilis] prf 2009358B Leu-rich protein [Streptococcus equisimilis] Length = 282
SeqID 988	SA-1916.2	Contig133 (7885-9018 m)	90	Identities = 320/377 (84%), Positives = 359/377 (94%) sp Q00752 MSMK_STRMU MULTIPLE SUGAR-BINDING TRANSPORT ATP-BINDING PROTEIN MSMK pir E42400 ABC-type transport system ATP-binding protein msmk [validated] - Streptococcus mutans gb AAA26938.1 (M77351) ATP-binding protein [Streptococcus mutans] Length = 377
SeqID 989	SA-1918.2	Contig82 (6068-7264 p)	91	Identities = 342/399 (85%), Positives = 378/399 (94%) emb CAC21180.1 (AJ251799) putative phosphopentomutase [Streptococcus thermophilus] Length = 403
SeqID 990	SA-1919.1	Contig82 (5330-6001 p)	57	Identities = 99/199 (49%), Positives = 140/199 (69%), Gaps = 1/199 (0%) sp P43048 DEOC_MYCHO DEOXYRIBOSE-PHOSPHATE ALDOLASE (PHOSPHODEOXYRIBOALDOLASE) (DEOXYRIBOALDOLASE) pir S42197 deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma hominis (SGC3) pir S72522 deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma hominis (SGC3) emb CAA81646.1 (Z27121) deoxyribose aldolase [Mycoplasma hominis] Length = 217
SeqID 991	SA-1921.1	Contig82 (4098-5300 p)	62	Identities = 160/405 (39%), Positives = 256/405 (62%), Gaps = 8/405 (1%) dbj BAB05165.1 (AP001512) nucleoside transporter [Bacillus halodurans] Length = 406

SeqID 992	SA-1922.1	Contig82 (3298-4077 p)	63	Identities = 145/246 (58%), Positives = 171/246 (68%) sp O83990 UDP_TREPA URIDINE PHOSPHORYLASE (UDRPASE) pir F71251 probable uridine phosphorylase (udp) - syphilis spirochete gb AAC65977.1 (AE001270) uridine phosphorylase (udp) [Treponema pallidum] Length = 258
SeqID 993	SA-1923.1	Contig82 (2403-3140 m)	33	Identities = 50/171 (29%), Positives = 86/171 (50%), Gaps = 17/171 (9%) dbj BAB06113.1 (AP001515) transcriptional regulator (GntR family) [Bacillus halodurans] Length = 242
SeqID 994	SA-1924.1	Contig82 (2094-2402 m)	No Hits found	
SeqID 995	SA-1925.1	Contig82 (803-1717 m)	45	Identities = 84/314 (26%), Positives = 140/314 (43%), Gaps = 28/314 (8%) pir S55315 mucin (clone PGM-2A) - pig pir I47141 gastric mucin (clone PGM-2A) - pig (fragment) gb AAC48526.1 (U10281) gastric mucin [Sus scrofa] Length = 528
SeqID 996	SA-1926.1	Contig82 (372-1994 p)	94	Identities = 471/539 (87%), Positives = 512/539 (94%), Gaps = 1/539 (0%) gb AAD23455.1 (AF117741) chaperonin GroEL [Streptococcus pneumoniae] Length = 540
SeqID 997	SA-1927.1	Contig82 (3-344 m)	No Hits found	
SeqID 998	SA-1928.1	Contig82 (16-276 p)	81	Identities = 52/91 (57%), Positives = 75/91 (82%), Gaps = 1/91 (1%) gb AAD23454.1 (AF117741) co-chaperonin GroES [Streptococcus pneumoniae] Length = 94
SeqID 999	SA-1929.2	Contig84 (8598-9938 m)	63	Identities = 191/454 (42%), Positives = 289/454 (63%), Gaps = 17/454 (3%) dbj BAB04579.1 (AP001510) BH0860-unknown conserved protein [Bacillus halodurans] Length = 458

SeqID 1000	SA-193.2	Contig136 (63475-63975 m)	64	Identities = 96/164 (58%), Positives = 125/164 (75%), Gaps = 1/164 (0%) sp P42923 RL10_BACSU 50S RIBOSOMAL PROTEIN L10 (BL5) (COLD ACCLIMATIZATION PROTEIN) (CAP) (VEGETATIVE PROTEIN 300) (VEG300) pir D69695 ribosomal protein L10 (BL5) rplJ - Bacillus subtilis dbj BAA08840.1 (D50303) Ribosomal Protein L10 [Bacillus subtilis] emb CAB11880.1 (Z99104) ribosomal protein L10 (BL5) [Bacillus subtilis] Length = 166
SeqID 1001	SA-1930.1	Contig84 (7772-8539 p)	44	Identities = 72/213 (33%), Positives = 115/213 (53%), Gaps = 11/213 (5%) pir C70361 conserved hypothetical protein aq_701 - Aquifex aeolicus gb AAC06891.1 (AE000703) hypothetical protein [Aquifex aeolicus] Length = 413
SeqID 1002	SA-1931.1	Contig84 (5832-7613 m)	73	Identities = 344/578 (59%), Positives = 446/578 (76%) sp P14951 UVR_C_BACSU EXCINUCLEASE ABC SUBUNIT C pir A37192 excinuclease ABC, chain C - Bacillus subtilis gb AAA87316.1 (J03294) deoxyribodipyrimidine photolyase [Bacillus subtilis] emb CAA99578.1 (Z75208) deoxyribodipyrimidine photolyase [Bacillus subtilis] emb CAB14809.1 (Z99118) excinuclease ABC (subunit C) [Bacillus subtilis] Length = 598
SeqID 1003	SA-1932.2	Contig84 (3955-5790 p)	45	Identities = 151/585 (25%), Positives = 281/585 (47%), Gaps = 31/585 (5%) gb AAK06095.1 AE006429_13 (AE006429) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 617
SeqID 1004	SA-1933.2	Contig84 (3237-3839 m)	65	Identities = 94/203 (46%), Positives = 134/203 (65%), Gaps = 3/203 (1%) gb AAK04278.1 AE006255_8 (AE006255) oxidoreductase [Lactococcus lactis subsp. lactis] Length = 203
SeqID 1005	SA-1934.2	Contig84 (1784-3190 m)	77	Identities = 311/474 (65%), Positives = 369/474 (77%), Gaps = 11/474 (2%) gb AAC45369.1 (U78036) dipeptidase [Lactococcus lactis] Length = 472

SeqID 1006	SA-1935.1	Contig84 (1103-1687 m)	76	Identities = 125/192 (65%), Positives = 163/192 (84%), Gaps = 1/192 (0%) sp O86222 Y22A_HAEN HYPOTHETICAL PROTEIN HI0220.2 gb AAC21888.1 (U32707) H. influenzae predicted coding region HI0220.2 [Haemophilus influenzae Rd] Length = 214
SeqID 1007	SA-1937.1	Contig84 (354-1088 m)	39	Identities = 70/244 (28%), Positives = 103/244 (41%), Gaps = 31/244 (12%) pir E72330 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD35886.1 AE001748_2 (AE001748) conserved hypothetical protein [Thermotoga maritima] Length = 233
SeqID 1008	SA-1939.1	Contig84 (3-185 m)	40	Identities = 27/63 (42%), Positives = 36/63 (56%) pir T31110 extracellular matrix binding protein - Abiotrophia defectiva (fragment) gb AAD03320.1 (AF067776) extracellular matrix binding protein [Abiotrophia defectiva] Length = 2055
SeqID 1009	SA-194.1	Contig136 (63046-63411 m)	73	Identities = 83/123 (67%), Positives = 96/123 (77%), Gaps = 2/123 (1%) sp P02394 RL7_BACSU 50S RIBOSOMAL PROTEIN L7/L12 (BL9) (A TYPE) (VEGETATIVE PROTEIN 341) (VEG341) pir R5BS9 ribosomal protein L7/L12 - Bacillus subtilis emb CAB11881.1 (Z99104) ribosomal protein L12 (BL9) [Bacillus subtilis] Length = 123
SeqID 1010	SA-1940.1	Contig103 (17991-18413 p)	85	Identities = 112/141 (79%), Positives = 124/141 (87%) sp Q9ZA56 HPRK_STRMU HPR(SER) KINASE/PHOSPHATASE pir T11568 probable HPr kinase (EC 2.7.1.-) - Streptococcus mutans gb AAC80172.1 (U75480) putative HPr(ser) kinase [Streptococcus mutans] Length = 311
SeqID 1011	SA-1941.1	Contig103 (17660-17875 p)	No Hits found	

SeqID 1012	SA-1942.1	Contig103 (17500-17763 p)	38	Identities = 21/63 (33%), Positives = 36/63 (56%), Gaps = 2/63 (3%) pir E70043 hypothetical protein yvC - Bacillus subtilis gb AAC67275.1 (AF017113) yvC [Bacillus subtilis] emb CAB15516.1 (Z99121) yvC [Bacillus subtilis] emb CAB15528.1 (Z99122) yvC [Bacillus subtilis] Length = 65
SeqID 1013	SA-1943.1	Contig103 (16961-17419 p)	57	Identities = 63/146 (43%), Positives = 92/146 (62%), Gaps = 5/146 (3%) dbj BAA90855.1 (AB031213) Ydck [Bacillus halodurans] dbj BAB04251.1 (AP001508) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 151
SeqID 1014	SA-1944.1	Contig103 (14836-16998 p)	67	Identities = 379/729 (51%), Positives = 515/729 (69%), Gaps = 25/729 (3%) dbj BAB04250.1 (AP001508) unknown conserved protein [Bacillus halodurans] Length = 728
SeqID 1015	SA-1945.1	Contig103 (13688-14731 p)	61	Identities = 164/300 (54%), Positives = 217/300 (71%), Gaps = 1/300 (0%) gb AAB94650.1 (U96107) N5.N10-methylenetetrahydromethanopterin reductase homolog [Staphylococcus carnosus] Length = 301
SeqID 1016	SA-1946.1	Contig103 (13427-13555 m)	No Hits found	
SeqID 1017	SA-1949.1	Contig103 (12519-13334 m)	64	Identities = 141/283 (49%), Positives = 198/283 (69%), Gaps = 3/283 (1%) pir E69759 hypothetical protein ycgR - Bacillus subtilis dbj BAA08959.1 (D50453) homologue of unidentified protein of E. coli [Bacillus subtilis] emb CAB12119.1 (Z99105) ycgR [Bacillus subtilis] Length = 294
SeqID 1018	SA-195.1	Contig136 (62894-63439 p)	45	Identities = 38/157 (24%), Positives = 82/157 (52%), Gaps = 6/157 (3%) pir T39903 serine-rich protein - fission yeast (Schizosaccharomyces pombe) emb CAA22127.1 (AL033534) hypothetical serine-rich secreted protein [Schizosaccharomyces pombe] Length = 534
SeqID 1019	SA-1950.1	Contig103 (12179-12337 p)	No Hits found	

SeqID 1020	SA-1951.1	Contig103 (11707-12519 m)	55	Identities = 101/290 (34%), Positives = 160/290 (54%), Gaps = 25/290 (8%) pir D69759 hypothetical protein ycgQ - Bacillus subtilis dbj BAA08958.1 (D50453) ycgQ [Bacillus subtilis] emb CAB12118.1 (Z99105) ycgQ [Bacillus subtilis] Length = 285
SeqID 1021	SA-1952.2	Contig103 (10060-11670 p)	61	Identities = 239/537 (44%), Positives = 330/537 (60%), Gaps = 79/537 (14%) gb AAK04911.1 AE006314_9 (AE006314) cell division protein FtsY [Lactococcus lactis subsp. lactis] Length = 459
SeqID 1022	SA-1953.1	Contig111 (18352-19497 m)	77	Identities = 231/381 (60%), Positives = 300/381 (78%), Gaps = 4/381 (1%) gb AAF91339.1 AF249729_1 (AF249729) ATPase OpuCA [Listeria monocytogenes] Length = 397
SeqID 1023	SA-1954.1	Contig111 (17717-18352 m)	74	Identities = 122/212 (57%), Positives = 162/212 (75%) gb AAF91340.1 AF249729_2 (AF249729) membrane permease OpuCB [Listeria monocytogenes] Length = 218
SeqID 1024	SA-1955.1	Contig111 (16788-17714 m)	62	Identities = 166/303 (54%), Positives = 222/303 (72%), Gaps = 1/303 (0%) sp O32243 OPCC_BACSU GLYCINE BETAINES/CARNITINE/CHOLINE-BINDING PROTEIN PRECURSOR (OSMOPROTECTANT-BINDING PROTEIN) pir E69670 glycine betaine/carnitine/choline ABC transporter (osmoprotect) opuCC - Bacillus subtilis emb CAB15386.1 (Z99121) glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein) [Bacillus subtilis] Length = 303
SeqID 1025	SA-1956.1	Contig111 (16137-16778 m)	82	Identities = 135/213 (63%), Positives = 179/213 (83%) gb AAF91342.1 AF249729_4 (AF249729) membrane permease OpuCD [Listeria monocytogenes] Length = 223

SeqID 1026	SA-1957.1	Contig111 (14894-15880 p)	50	Identities = 85/335 (25%), Positives = 171/335 (50%), Gaps = 15/335 (4%) pir H72265 hypothetical protein TM1336 - Thermotoga maritima (strain MSB8) gb AAD36408.1 AE001788_3 (AE001788) permease, putative [Thermotoga maritima] Length = 390
SeqID 1027	SA-1958.1	Contig111 (13949-14821 p)	49	Identities = 75/279 (26%), Positives = 144/279 (50%) sp P49330 RGG_STRGC_RGG_PROTEIN pir A41898 positive regulator of glucosyltransferase and Spp phenotype rgg - Streptococcus gordonii gb AA26968.1 (M89776) rgg [Streptococcus gordonii] Length = 297
SeqID 1028	SA-1959.1	Contig111 (12837-13118 p)	No Hits found	
SeqID 1029	SA-196.1	Contig136 (62632-62808 m)	No Hits found	
SeqID 1030	SA-1960.1	Contig111 (12417-12803 p)	No Hits found	
SeqID 1031	SA-1961.1	Contig111 (12013-12432 p)	No Hits found	
SeqID 1032	SA-1962.2	Contig111 (11221-11808 p)	24	Identities = 32/126 (25%), Positives = 52/126 (40%), Gaps = 20/126 (15%) pir F71614 chromatinic RING finger DRING protein homolog PFB0440c - malaria parasite (Plasmodium falciparum) gb AAC71877.1 (AE001395) chromatinic RING finger protein, DRING ortholog [Plasmodium falciparum] Length = 568
SeqID 1033	SA-1963.2	Contig111 (10793-11221 p)	37	Identities = 33/113 (29%), Positives = 53/113 (46%), Gaps = 19/113 (16%) pir E71619 RAD2 endonuclease PFB0265c - malaria parasite (Plasmodium falciparum) gb AAC71842.1 (AE001383) RAD2 endonuclease [Plasmodium falciparum] Length = 1516
SeqID 1034	SA-1966.1	Contig136 (26286-28304 m)	43	Identities = 155/698 (22%), Positives = 293/698 (41%), Gaps = 157/698 (22%) pir F70175 rep helicase, single-stranded DNA-dependent ATPase (rep) homolog - Lyme disease spirochete gb AAC66967.1 (AE001162) rep helicase, single-stranded DNA-dependent ATPase (rep) [Borrelia burgdorferi] Length = 659

SeqID 1035	SA-1967.1	Contig136 (25592-25951 m)	55	Identities = 61/118 (51%), Positives = 79/118 (66%), Gaps = 1/118 (0%) gb AAC98436.1 (L29324) unknown [Streptococcus pneumoniae] Length = 118
SeqID 1036	SA-1968.1	Contig136 (25217-25582 m)	76	Identities = 72/121 (59%), Positives = 95/121 (78%) gb AAC98435.1 (L29324) unknown [Streptococcus pneumoniae] Length = 121
SeqID 1037	SA-1969.1	Contig136 (23353-25230 m)	49	Identities = 241/416 (57%), Positives = 312/416 (74%), Gaps = 4/416 (0%) gb AAC98434.1 (L29324) relaxase [Streptococcus pneumoniae] Length = 431
SeqID 1038	SA-197.1	Contig136 (61820-62635 m)	81	Identities = 195/268 (72%), Positives = 233/268 (86%), Gaps = 1/268 (0%) gb AAD41248.1 AF106927_1 (AF106927) unknown [Streptococcus suis] Length = 272
SeqID 1039	SA-1971.2	Contig136 (22443-23198 m)	77	Identities = 146/250 (58%), Positives = 197/250 (78%) sp P26422 LACR_STRMU LACTOSE PHOSPHOTRANSFERASE SYSTEM REPRESSOR pir B43258 regulatory protein lacR - Streptococcus mutans gb AAA26903.1 (M80797) lactose repressor [Streptococcus mutans] Length = 251
SeqID 1040	SA-1973.3	Contig117 (6819-7409 m)	54	Identities = 78/193 (40%), Positives = 111/193 (57%), Gaps = 14/193 (7%) gb AAK04415.1 AE006268_11 (AE006268) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 191
SeqID 1041	SA-1974.2	Contig117 (5976-6734 p)	No Hits found	
SeqID 1042	SA-1975.2	Contig117 (5755-5973 p)	59	Identities = 28/60 (46%), Positives = 45/60 (74%) pir G75183 probable repressor protein PAB7155 - Pyrococcus abyssi (strain Orsay) emb CAB49550.1 (AJ248284) repressor protein, putative [Pyrococcus abyssi] Length = 73
SeqID 1043	SA-1976.2	Contig117 (5390-5713 p)	No Hits found	

SeqID 1044	SA-1977.1	Contig117 (4126-4944 p)	42	<p>Identities = 70/258 (27%), Positives = 122/258 (47%), Gaps = 14/258 (5%) sp P27129 RF AJ_ECOLI LIPOPOLYSACCHARIDE 1,2-GLUCOSYLTRANSFERASE pir S47847 lipopolysaccharide glucosyltransferase I (EC 2.4.1.58) - Escherichia coli gb AAB18603.1 (U00039) UDP-D-glucose:(galactosyl)lipopolysaccharide glucosyltransferase [Escherichia coli] gb AAC76650.1 (AE000440) UDP-D-glucose:(galactosyl)lipopolysaccharide glucosyltransferase [Escherichia coli K12] Length = 338</p>
SeqID 1045	SA-1978.1	Contig117 (3322-4038 p)	56	<p>Identities = 84/242 (34%), Positives = 147/242 (60%), Gaps = 16/242 (6%) dbj BAB07774.1 (AP001520) unknown conserved protein [Bacillus halodurans] Length = 236</p>
SeqID 1046	SA-198.1	Contig136 (61161-61589 m)	71	<p>Identities = 90/141 (63%), Positives = 102/141 (71%) gb AAC98422.1 (L29323) methyl transferase [Streptococcus pneumoniae] Length = 172</p>
SeqID 1047	SA-1982.1	Contig117 (622-3240 p)	70	<p>Identities = 482/885 (54%), Positives = 618/885 (69%), Gaps = 27/885 (3%) dbj BAB04986.1 (AP001511) alanyl-tRNA synthetase [Bacillus halodurans] Length = 879</p>
SeqID 1048	SA-1983.1	Contig117 (121-606 p)	42	<p>Identities = 43/144 (29%), Positives = 75/144 (51%), Gaps = 2/144 (1%) pir G70079 hypothetical protein yxj - Bacillus subtilis dbj BAA11710.1 (D83026) hypothetical [Bacillus subtilis] emb CAB15920.1 (Z99123) yxj [Bacillus subtilis] Length = 162</p>
SeqID 1049	SA-1985.2	Contig128 (12351-12893 m)	55	<p>Identities = 60/175 (34%), Positives = 103/175 (58%), Gaps = 6/175 (3%) gb AAK05744.1 (AE006395_4) (AE006395) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 186</p>

SeqID 1050	SA-1986.1	Contig128 (9927-12275 m)	58	Identities = 321/790 (40%), Positives = 466/790 (58%), Gaps = 18/790 (2%) sp P94545 MUS2_BACSU MUTS2 PROTEIN pir D6985 DNA mismatch repair protein homolog yshD - Bacillus subtilis emb CAA99569.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14818.1 (Z99118) similar to DNA mismatch repair protein [Bacillus subtilis] Length = 785
SeqID 1051	SA-1987.1	Contig128 (9423-9770 m)	42	Identities = 42/143 (29%), Positives = 71/143 (49%), Gaps = 7/143 (4%) sp Q57819 Y374_METJA HYPOTHETICAL PROTEIN MJ0374 pir F64346 hypothetical protein MJ0374 - Methanococcus jannaschii gb AAB98363.1 (U67490) lipoprotein B (lppB) [Methanococcus jannaschii] Length = 330
SeqID 1052	SA-1988.1	Contig128 (9028-9360 m)	60	Identities = 64/100 (64%), Positives = 78/100 (78%), Gaps = 1/100 (1%) emb CAB40815.2 (AJ133006) thioredoxin [Listeria monocytogenes] Length = 103
SeqID 1053	SA-1989.1	Contig128 (8389-8982 p)	25	Identities = 27/67 (40%), Positives = 51/67 (75%) pir 40868 hypothetical protein 3 nanH region - Clostridium perfringens emb CAA60798.1 (X87369) ORF3 [Clostridium perfringens] Length = 265
SeqID 1054	SA-199.1	Contig136 (60382-61734 m)	85	Identities = 322/448 (71%), Positives = 386/448 (85%), Gaps = 3/448 (0%) gb AAC98421.1 (L29323) methyl transferase [Streptococcus pneumoniae] Length = 452
SeqID 1055	SA-1990.2	Contig128 (7088-8212 p)	58	Identities = 169/388 (43%), Positives = 224/388 (57%), Gaps = 26/388 (6%) gb AAK04939.1 AE006318_2 (AE006318) A/G-specific adenine glycosylase (EC 3.2.2.1) [Lactococcus lactis subsp. lactis] Length = 387
SeqID 1056	SA-1992.1	Contig128 (47161-47817 m)	19	Identities = 31/105 (29%), Positives = 51/105 (48%), Gaps = 4/105 (3%) emb CAB95931.1 (AL359989) galactose-1-phosphate uridylyltransferase [Streptomyces coelicolor A3(2)] Length = 353

SeqID 1057	SA-1993.1	Contig128 (45990-47117 p)	65	Identities = 186/410 (45%), Positives = 259/410 (62%), Gaps = 27/410 (6%) dbj BAB06998.1 (AP001518) unknown conserved protein [Bacillus halodurans] Length = 422
SeqID 1058	SA-1994.1	Contig128 (45274-45810 m)	44	Identities = 51/173 (29%), Positives = 87/173 (49%), Gaps = 20/173 (11%) emb CAA73267.1 (Y12736) orfX [Lactococcus lactis subsp. cremoris] Length = 200
SeqID 1059	SA-1995.1	Contig128 (44643-45251 m)	24	Identities = 29/87 (33%), Positives = 51/87 (58%), Gaps = 1/87 (1%) dbj BAB06137.1 (AP001515) DNA polymerase III (alpha subunit) [Bacillus halodurans] Length = 1433
SeqID 1060	SA-1996.1	Contig128 (43791-44516 m)	29	Identities = 46/141 (32%), Positives = 71/141 (49%), Gaps = 11/141 (7%) sp P32184 TIPA_STRCO TRANSCRIPTIONAL ACTIVATOR TIPA pir S35354 tipA protein - Streptomyces lividans pir T36339 transcription regulator - Streptomyces coelicolor gb AAB27737.1 (S64314) TipAL-AS=thiostrepton-specific recognition protein(TipAL=transcriptional activator, TipAS=transcriptional activation modulator) [Streptomyces lividans, Peptide, 253 aa] emb CAB42766.1 (AL049841) transcriptional regulator [Streptomyces coelicolor A3(2)] Length = 253
SeqID 1061	SA-1997.1	Contig128 (42531-43721 p)	63	Identities = 166/397 (41%), Positives = 255/397 (63%), Gaps = 14/397 (3%) pir H64571 cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79) - Helicobacter pylori (strain 26695) gb AAD07482.1 (AE000557) cyclopropane fatty acid synthase (cfa) [Helicobacter pylori 26695] Length = 389

SeqID 1062	SA-1998.1	Contig128 (42071-42370 p)	44	Identities = 44/97 (45%), Positives = 60/97 (61%) sp P24281 YAAK_BACSU HYPOTHETICAL 11.8 KD PROTEIN IN DNAZ-RECR INTERGENIC REGION pir S13787 conserved hypothetical protein yaak - Bacillus subtilis emb CAA34878.1 (X17014) ORF107 [Bacillus subtilis] dbj BAA05256.1 (D26185) unknown [Bacillus subtilis] emb CAB11796.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 107
SeqID 1063	SA-1999.1	Contig128 (41421-41759 m)	No Hits found	
SeqID 1064	SA-2.1	Contig137 (42467-42724 p)	No Hits found	
SeqID 1065	SA-20.1	Contig137 (24860-25942 p)	20	Identities = 37/161 (22%), Positives = 75/161 (45%), Gaps = 8/161 (4%) pir S45089 hypothetical protein eta - Streptococcus pyogenes plasmid pDB101 pir S68123 hypothetical protein eta - Streptococcus pyogenes plasmid pMD101 emb CAA47095.1 (X66468) orf eta [Streptococcus pyogenes] Length = 231
SeqID 1066	SA-200.1	Contig136 (59952-60398 m)	59	Identities = 69/143 (48%), Positives = 97/143 (67%), Gaps = 3/143 (2%) gb AAC98423.1 (L29323) unknown [Streptococcus pneumoniae] Length = 149
SeqID 1067	SA-2000.1	Contig128 (40742-41188 m)	No Hits found	
SeqID 1068	SA-2004.1	Contig128 (151-2997 m)	82	Identities = 640/941 (68%), Positives = 784/941 (83%), Gaps = 3/941 (0%) sp O34863 UVRA_BACSU EXCINUCLEASE ABC SUBUNIT A pir F69729 excinuclease ABC chain A - Bacillus subtilis gb AAC67271.1 (AF017113) excinuclease ABC subunit A [Bacillus subtilis] emb CAB15533.1 (Z99122) excinuclease ABC (subunit A) [Bacillus subtilis] Length = 957
SeqID 1069	SA-2005.1	Contig128 (3111-3782 m)	No Hits found	
SeqID 1070	SA-2006.1	Contig128 (3807-4751 m)	50	Identities = 108/295 (36%), Positives = 174/295 (58%), Gaps = 4/295 (1%) pir A75272 probable transport protein - Deinococcus radiodurans (strain R1) gb AAF12002.1 AE002075_6 (AE002075) transport protein, putative [Deinococcus radiodurans] Length = 312

SeqID 1071	SA-2007.1	Contig128 (4921-5160 m)	73	Identities = 52/79 (65%), Positives = 64/79 (80%) ribosomal protein S18 (rpsR) - Bacillus subtilis dbjBAA05219.1 (D26185) ribosomal protein S18 [Bacillus subtilis] emb CAB16126.1 (Z99124) ribosomal protein S18 [Bacillus subtilis] Length = 81
SeqID 1072	SA-2008.2	Contig128 (5205-5696 m)	90	Identities = 136/163 (83%), Positives = 149/163 (90%) gb AAF98351.1 (AF280767) single strand binding protein [Streptococcus pyogenes] Length = 163
SeqID 1073	SA-2009.2	Contig128 (5708-5995 m)	78	Identities = 72/96 (75%), Positives = 81/96 (84%), Gaps = 1/96 (1%) gb AAK06289.1 AE006448_7 (AE006448) 30S ribosomal protein S6 [Lactococcus lactis subsp. lactis] Length = 97
SeqID 1074	SA-2010.1	Contig113 (25825-26142 p)	64	Identities = 52/93 (55%), Positives = 68/93 (72%) gb AAC14608.1 (U95840) transmembrane protein Tmp5 [Lactococcus lactis] Length = 273
SeqID 1075	SA-2012.1	Contig113 (25470-25775 m)	No Hits found	
SeqID 1076	SA-2013.1	Contig113 (22823-23569 m)	40	Identities = 92/204 (45%), Positives = 101/204 (49%), Gaps = 1/204 (0%) pir H72504 hypothetical protein APE2014 - Aeropyrum pernix (strain K1) dbj BAA81024.1 (AP000063) 280aa long hypothetical protein [Aeropyrum pernix] Length = 280
SeqID 1077	SA-2014.1	Contig113 (22538-22783 m)	No Hits found	
SeqID 1078	SA-2015.2	Contig113 (22413-24389 p)	85	Identities = 490/652 (75%), Positives = 561/652 (85%), Gaps = 5/652 (0%) gb AAC16243.2 (AF061748) cell division protein FtsH [Streptococcus pneumoniae] Length = 652

SeqID 1079	SA-2017.2	Contig113 (21848-22390 p)	81	<p>Identities = 121/179 (67%), Positives = 152/179 (84%), Gaps = 1/179 (0%) sp Q02522 HPRT_LACLA HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE (HGPR) (HGPRASE) pir S30100 hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - Lactococcus lactis emb CAA48876.1 (X69123) hypoxanthine guanine phosphoribosyltransferase [Lactococcus lactis] emb CAA47404.1 (X67015) hypoxanthine phosphoribosyltransferase [Lactococcus lactis] gb AAK04118.1 AE006241_7 (AE006241) hypoxanthine-guanine phosphorybosyltransferase (EC 2.4.2.8) [Lactococcus lactis subsp. lactis] prf 1905381A hypoxanthine guanine phosphoribosyltransferase [Lactococcus lactis] Length = 183</p>
SeqID 1080	SA-2018.2	Contig113 (20569-21843 p)	53	<p>Identities = 142/418 (33%), Positives = 230/418 (54%), Gaps = 21/418 (5%) gb AAK04117.1 AE006241_6 (AE006241) cell cycle protein MesJ [Lactococcus lactis subsp. lactis] Length = 423</p>
SeqID 1081	SA-2019.2	Contig113 (19401-20567 p)	26	<p>Identities = 77/284 (27%), Positives = 117/284 (41%), Gaps = 54/284 (19%) prf 1714238A beta lactamase mutant S-3P [Staphylococcus aureus] Length = 281</p>
SeqID 1082	SA-202.1	Contig136 (59573-59959 m)	No Hits found	
SeqID 1083	SA-2022.2	Contig93 (6702-7730 p)	68	<p>Identities = 159/334 (47%), Positives = 236/334 (70%), Gaps = 15/334 (4%) gb AAK05774.1 AE006398_5 (AE006398) transcriptional regulator [Lactococcus lactis subsp. lactis] Length = 324</p>
SeqID 1084	SA-2024.1	Contig93 (5199-6446 m)	42	<p>Identities = 121/418 (28%), Positives = 190/418 (44%), Gaps = 43/418 (10%) sp P29850 MALX_STRPN MALTOSE/MALTODEXTRIN-BINDING PROTEIN PRECURSOR pir S32569 malX protein - Streptococcus pneumoniae gb AAA26925.1 (L08611) MalX [Streptococcus pneumoniae] Length = 423</p>

SeqID 1085	SA-2025.1	Contig93 (3731-5071 m)	56	Identities = 174/428 (40%), Positives = 263/428 (60%), Gaps = 21/428 (4%) pir S63615 malF protein homolog cymF - Klebsiella oxytoca emb CAA60005.1 (X86014) cymF [Klebsiella oxytoca] Length = 427
SeqID 1086	SA-2026.1	Contig93 (2895-3731 m)	61	Identities = 109/269 (40%), Positives = 171/269 (63%), Gaps = 5/269 (1%) pir S63616 malG protein homolog cymG - Klebsiella oxytoca emb CAA60006.1 (X86014) cymG [Klebsiella oxytoca] Length = 277
SeqID 1087	SA-2028.1	Contig93 (1192-2574 m)	69	Identities = 259/485 (53%), Positives = 353/485 (72%), Gaps = 9/485 (1%) sp P94408 YCLF_BACSU HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION pir C69762 di-tripeptide ABC transporter (membrane pr) homolog yclF - Bacillus subtilis dbj BAA09000.1 (D50453) homolog of Di-tripeptide transporter Dtp of L. lactis [Bacillus subtilis] emb CAB12175.1 (Z99106) similar to di-tripeptide ABC transporter (membrane protein) [Bacillus subtilis] Length = 492
SeqID 1088	SA-2029.1	Contig93 (683-1147 p)	40	Identities = 37/114 (32%), Positives = 66/114 (57%), Gaps = 10/114 (8%) dbj BAB07289.1 (AP001519) mutator MutT protein [Bacillus halodurans] Length = 159
SeqID 1089	SA-203.1	Contig136 (59343-59546 m)	No Hits found	
SeqID 1090	SA-2030.1	Contig93 (411-599 p)	No Hits found	
SeqID 1091	SA-2031.1	Contig93 (20-400 p)	52	Identities = 41/117 (35%), Positives = 67/117 (57%) gb AAAF61315.1 (U96166) unknown [Streptococcus cristatus] Length = 442
SeqID 1092	SA-2033.2	Contig123 (38769-39380 p)	52	Identities = 73/188 (38%), Positives = 108/188 (56%), Gaps = 11/188 (5%) dbj BAA21095.1 (D88438) repB [Lactobacillus acidophilus] Length = 193
SeqID 1093	SA-2034.1	Contig123 (37952-38569 p)	No Hits found	
SeqID 1094	SA-2035.2	Contig123 (36936-37967 p)	22	Identities = 50/196 (25%), Positives = 85/196 (42%), Gaps = 12/196 (6%) gb AAB60012.1 (U09422) ORF21 [Enterococcus faecalis] prf 2114402D ORF 21 [Enterococcus faecalis] Length = 461

SeqID 1095	SA-2037.2	Contig123 (36547-36867 p)	No Hits found	
SeqID 1096	SA-2038.1	Contig123 (35896-36345 p)	No Hits found	
SeqID 1097	SA-2040.1	Contig123 (34160-35575 p)	No Hits found	
SeqID 1098	SA-2041.2	Contig123 (33771-34109 p)	60	Identities = 50/110 (45%), Positives = 76/110 (68%) ref NP_054018.1 CadX [Staphylococcus lugdunensis] gb AAB18271.1 (U74623) CadX [Staphylococcus lugdunensis] Length = 115
SeqID 1099	SA-2042.2	Contig123 (33133-33759 p)	97	Identities = 198/209 (94%), Positives = 203/209 (96%) pir E81967 cadmium resistance protein NMA0496 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB83789.1 (AL162753) cadmium resistance protein [Neisseria meningitidis Z2491] Length = 213
SeqID 1100	SA-2045.2	Contig132 (41522-42709 p)	59	Identities = 171/393 (43%), Positives = 256/393 (64%), Gaps = 2/393 (0%) gb AAK06200.1 AE006438_14 (AE006438) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 400
SeqID 1101	SA-2047.1	Contig132 (39834-41324 p)	51	Identities = 186/500 (37%), Positives = 258/500 (51%), Gaps = 67/500 (13%) emb CAB95221.1 (AL359773) possible threonine synthase [Leishmania major] Length = 676
SeqID 1102	SA-2048.1	Contig132 (38697-39713 p)	76	Identities = 238/340 (70%), Positives = 276/340 (81%), Gaps = 2/340 (0%) pir H81186 alcohol dehydrogenase, propanol- preferring NMB0546 [imported] - Neisseria meningitidis (group B strain MD58) gb AAF40975.1 (AE002410) alcohol dehydrogenase, propanol-preferring [Neisseria meningitidis MC58] Length = 348
SeqID 1103	SA-205.1	Contig136 (58759-59349 m)	No Hits found	
SeqID 1104	SA-2050.2	Contig132 (35876-38518 p)	84	Identities = 658/873 (75%), Positives = 760/873 (86%), Gaps = 2/873 (0%) gb AAK03537.1 (AE006181) Adh2 [Pasteurella multocida] Length = 875
SeqID 1105	SA-2051.1	Contig129 (2-493 m)	58	Identities = 68/155 (43%), Positives = 98/155 (62%) gb AAF13747.1 AF117351_4 (AF117351) unknown [Zymomonas mobilis] Length = 236

SeqID 1106	SA-2052.1	Contig129 (646-1674 p)	80	Identities = 227/341 (66%), Positives = 278/341 (80%) sp O32054 QUEA_BACSU S-ADENOSYLMETHIONINE:TRNA RIBOSYLTRANSFERASE-ISOMERASE (QUEUOSINE BIOSYNTHESIS PROTEIN QUEA) pir A69688 S- adenosylmethionine--tRNA ribosyltransferase-isomerase (EC 5.4.99.-) queA [similarity] - Bacillus subtilis emb CAB14732.1 (Z99118) S-adenosylmethionine tRNA ribosyltransferase [Bacillus subtilis] emb CAB75332.1 (Y15896) QueA protein [Bacillus subtilis] Length = 342
SeqID 1107	SA-2053.1	Contig129 (1761-2198 p)	54	Identities = 56/145 (38%), Positives = 86/145 (58%), Gaps = 2/145 (1%) emb CAA73494.1 (Y13052) ORF145 [Staphylococcus sciuri] Length = 145
SeqID 1108	SA-2054.1	Contig129 (2253-3317 m)	No Hits found	
SeqID 1109	SA-2056.1	Contig129 (3418-4674 m)	74	Identities = 232/416 (55%), Positives = 314/416 (74%), Gaps = 3/416 (0%) sp Q57493 Y092_HAEIN HYPOTHETICAL PROTEIN HI0092 pir D64142 hypothetical protein HI0092 - Haemophilus influenzae (strain Rd KW20) gb AAC21770.1 (U32694) H. influenzae predicted coding region HI0092 [Haemophilus influenzae Rd] Length = 419
SeqID 1110	SA-2059.2	Contig129 (4699-5841 m)	59	Identities = 177/367 (48%), Positives = 235/367 (63%), Gaps = 2/367 (0%) gb AAG58254.1 AE005541_6 (AE005541) orf, hypothetical protein [Escherichia coli O157:H7] Length = 387
SeqID 1111	SA-206.1	Contig136 (58190-58678 m)	No Hits found	
SeqID 1112	SA-2060.1	Contig129 (6008-7102 p)	50	Identities = 94/370 (25%), Positives = 184/370 (49%), Gaps = 13/370 (3%) dbj BAB06450.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 371
SeqID 1113	SA-2061.2	Contig129 (7171-8598 m)	80	Identities = 334/475 (70%), Positives = 392/475 (82%), Gaps = 8/475 (1%) gb AAAF89979.1 AF206272_5 (AF206272) beta- glucosidase [Streptococcus mutans] Length = 479

SeqID 1114	SA-2062.2	Contig132 (47810-48070 p)	95	<p>Identities = 82/86 (95%), Positives = 83/86 (96%)</p> <p>-sp Q9VWW03 RS17_STRPN 30S RIBOSOMAL PROTEIN S17- gb AAD33265.1 AF126059_6 (AF126059) RpS17 [Streptococcus pneumoniae] gb AAD33274.1 (AF126060) RpS17 [Streptococcus pneumoniae] gb AAD33283.1 (AF126061) RpS17 [Streptococcus pneumoniae] Length = 86</p>
SeqID 1115	SA-2063.2	Contig132 (47578-47784 p)	88	<p>Identities = 58/68 (85%), Positives = 64/68 (93%)</p> <p>sp Q9VWW08 RL29_STRPN 50S RIBOSOMAL PROTEIN L29 gb AAD33264.1 AF126059_5 (AF126059) RpL29 [Streptococcus pneumoniae] gb AAD33273.1 (AF126060) RpL29 [Streptococcus pneumoniae] gb AAD33282.1 (AF126061) RpL29 [Streptococcus pneumoniae] Length = 68</p>
SeqID 1116	SA-2065.1	Contig132 (47155-47568 p)	98	<p>Identities = 135/137 (98%), Positives = 137/137 (99%)</p> <p>sp Q9X5K1 RL16_STRPN 50S RIBOSOMAL PROTEIN L16 gb AAD33263.1 AF126059_4 (AF126059) RpL16 [Streptococcus pneumoniae] Length = 137</p>
SeqID 1117	SA-2066.1	Contig132 (46498-47151 p)	91	<p>Identities = 200/208 (96%), Positives = 203/208 (97%)</p> <p>sp Q9VWW37 RS3_STRPN 30S RIBOSOMAL PROTEIN S3 gb AAD33262.1 AF126059_3 (AF126059) RpS3 [Streptococcus pneumoniae] gb AAD33271.1 (AF126060) RpS3 [Streptococcus pneumoniae] gb AAD33280.1 (AF126061) RpS3 [Streptococcus pneumoniae] Length = 208</p>
SeqID 1118	SA-2067.1	Contig132 (46141-46485 p)	89	<p>Identities = 99/114 (86%), Positives = 106/114 (92%)</p> <p>sp Q9VWU5 RL22_STRPN 50S RIBOSOMAL PROTEIN L22 gb AAD33261.1 AF126059_2 (AF126059) RpL22 [Streptococcus pneumoniae] gb AAD33270.1 (AF126060) RpL22 [Streptococcus pneumoniae] gb AAD33279.1 (AF126061) RpL22 [Streptococcus pneumoniae] Length = 114</p>

SeqID 1119	SA-2069.1	Contig132 (45847-46125 p)	86	<p>Identities = 92/93 (98%), Positives = 93/93 (99%) sp Q9WW12 RS19_STRPN-30S RIBOSOMAL PROTEIN S19 gb AAD33260.1 AF126059.1 (AF126059) RpS19 [Streptococcus pneumoniae] gb AAD33269.1 (AF126060) RpS19 [Streptococcus pneumoniae] gb AAD33278.1 (AF126061) RpS19 [Streptococcus pneumoniae] Length = 93</p>
SeqID 1120	SA-207.1	Contig136 (56373-58190 m)	47	<p>Identities = 183/492 (37%), Positives = 292/492 (59%), Gaps = 30/492 (6%) gb AAF72343.1 AF192329_4 (AF192329) TrsK-like protein [Enterococcus faecalis] Length = 564</p>
SeqID 1121	SA-2071.1	Contig132 (44915-45748 p)	84	<p>Identities = 207/277 (74%), Positives = 239/277 (85%) gb AAC45959.1 (U43929) L2 [Bacillus subtilis] Length = 277</p>
SeqID 1122	SA-2072.1	Contig132 (44601-44897 p)	60	<p>Identities = 56/92 (60%), Positives = 67/92 (71%), Gaps = 1/92 (1%) sp Q9Z9L2 RL23_BACHD 50S RIBOSOMAL PROTEIN L23 pir T44385 ribosomal protein L23 [imported] - Bacillus halodurans dbj BAA75273.1 (AB017508) rplW homologue (identity of 71 to B. subtilis%) [Bacillus halodurans] dbj BAB03855.1 (AP001507) ribosomal protein L23 [Bacillus halodurans] Length = 96</p>
SeqID 1123	SA-2073.1	Contig132 (43978-44601 p)	73	<p>Identities = 130/207 (62%), Positives = 160/207 (76%) sp P42921 RL4_BACSU 50S RIBOSOMAL PROTEIN L4 pir H69694 ribosomal protein L4 rplD - Bacillus subtilis dbj BAA08832.1 (D50302) Ribosomal Protein L4 [Bacillus subtilis] gb AAC45957.1 (U43929) L4 [Bacillus subtilis] emb CAB11893.1 (Z99104) ribosomal protein L4 [Bacillus subtilis] Length = 207</p>
SeqID 1124	SA-2074.2	Contig132 (43328-43954 p)	82	<p>Identities = 157/208 (75%), Positives = 180/208 (86%), Gaps = 2/208 (0%) sp P42920 RL3_BACSU 50S RIBOSOMAL PROTEIN L3 (BL3) pir G69694 ribosomal protein L3 (BL3) rplC - Bacillus subtilis gb AAC45956.1 (U43929) L3 [Bacillus subtilis] emb CAB11892.1 (Z99104) ribosomal protein L3 (BL3) [Bacillus subtilis] Length = 209</p>

SeqID 1125	SA-2075.2	Contig132 (42915-43223 p)	94	Identities = 98/102 (96%), Positives = 102/102 (99%) sp P48853 RS10_STRMU_30S_RIBOSOMAL_PROTEIN_S10 gb AAB46363.1 (L29637) S10 ribosomal protein [Streptococcus mutans] Length = 102
SeqID 1126	SA-2076.2	Contig132 (42879-43280 m)	No Hits found	
SeqID 1127	SA-2077.1	Contig92 (5603-6382 p)	84	Identities = 241/259 (93%), Positives = 248/259 (95%) emb CAB90834.1 (AJ250837) putative transposase [Streptococcus dysgalactiae] Length = 259
SeqID 1128	SA-2078.1	Contig92 (5277-5567 p)	86	Identities = 93/96 (96%), Positives = 94/96 (97%) emb CAB90833.1 (AJ250837) hypothetical protein [Streptococcus dysgalactiae] Length = 96
SeqID 1129	SA-2079.1	Contig92 (4012-5049 m)	No Hits found	
SeqID 1130	SA-208.1	Contig136 (56111-56353 m)	No Hits found	
SeqID 1131	SA-2082.1	Contig92 (24-3476 m)	98	Identities = 114/1150 (99%), Positives = 1142/1150 (99%) gb AAB17762.1 (U56908) SCPB [Streptococcus agalactiae] Length = 1150
SeqID 1132	SA-2083.2	Contig102 (16409-17926 m)	74	Identities = 322/472 (68%), Positives = 378/472 (79%) pir S68598 sucrose-6-phosphate hydrolase ScrB - Streptococcus sobrinus (strain 6715) Length = 479
SeqID 1133	SA-2084.1	Contig102 (15445-16407 m)	84	Identities = 225/320 (70%), Positives = 273/320 (85%) sp Q54430 SCRR_STRMU_SUCROSE_OPERON_REPRESSOR (SCR_OPERON_REGULATORY_PROTEIN) gb AAC31628.1 (U46902) ScrR [Streptococcus mutans] Length = 320
SeqID 1134	SA-2085.1	Contig102 (14924-15358 p)	56	Identities = 51/129 (39%), Positives = 82/129 (63%), Gaps = 9/129 (6%) sp P54520 NUSB_BACSU_N_UTILIZATION SUBSTANCE_PROTEIN_B_HOMOLOG (NUSB_PROTEIN) pir F69960 transcription termination factor nusB homolog yqhZ [similarity] - Bacillus subtilis dbj BAA12571.1 (D84432) YqhZ [Bacillus subtilis] emb CAB14363.1 (Z99116) similar to transcription termination [Bacillus subtilis] Length = 131

SeqID 1135	SA-2086.1	- Contig102 (14542-14931 p)	53	Identities = 42/107 (39%), Positives = 70/107 (65%), Gaps = 4/107 (3%) dbj BAB06505.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 132
SeqID 1136	SA-2087.1	Contig102 (13893-14453 p)	64	Identities = 89/186 (47%), Positives = 120/186 (63%), Gaps = 1/186 (0%) sp P49778 EFP_BACSU ELONGATION FACTOR P (EF-P) pir JA69620 translation elongation factor EF-P efp - Bacillus subtilis dbj BAA12558.1 (D84432) YqhU [Bacillus subtilis] emb CAB14376.1 (Z99116) elongation factor P [Bacillus subtilis] Length = 185
SeqID 1137	SA-2090.1	Contig102 (10993-12714 p)	52	Identities = 174/568 (30%), Positives = 300/568 (52%), Gaps = 7/568 (1%) gb AAD10394.1 (U46488) NrpB [Proteus mirabilis] Length = 575
SeqID 1138	SA-2091.1	Contig102 (10920-11162 m)	No Hits found	
SeqID 1139	SA-2092.2	Contig102 (9291-11003 p)	50	Identities = 186/583 (31%), Positives = 305/583 (51%), Gaps = 14/583 (2%) gb AAD10393.1 (U46488) NrpA [Proteus mirabilis] Length = 588
SeqID 1140	SA-2095.2	Contig119 (7199-8548 m)	90	Identities = 377/449 (83%), Positives = 414/449 (91%) pir T51720 glucose-6-phosphate isomerase (EC 5.3.1.9) [imported] - Streptococcus mutans gb AAD33517.1 AF132127_2 (AF132127) glucose-6-phosphate isomerase [Streptococcus mutans] Length = 449
SeqID 1141	SA-2096.1	Contig119 (6350-6877 m)	71	Identities = 96/173 (55%), Positives = 129/173 (74%) dbj BAA28715.1 (AB001562) hypothetical protein [Streptococcus mutans] Length = 178
SeqID 1142	SA-2097.1	Contig119 (5682-6359 m)	72	Identities = 126/218 (57%), Positives = 166/218 (75%) emb CAB90755.1 (AJ400707) hypothetical protein [Streptococcus uberis] Length = 223

SeqID 1143	SA-2099.1	Contig119 (4507-5550 m)	60	Identities = 148/349 (42%), Positives = 223/349 (63%), Gaps = 16/349 (4%) sp O05252 YUFN_BACSU HYPOTHETICAL LIPOPROTEIN YUFN PRECURSOR pir C70009 ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis emb CAB07936.1 (Z93937) unknown [Bacillus subtilis] emb CAB15143.1 (Z99120) similar to ABC transporter (lipoprotein) [Bacillus subtilis] Length = 350
SeqID 1144	SA-21.1	Contig137 (24502-24849 p)	No Hits found	
SeqID 1145	SA-210.1	Contig136 (55240-56094 m)	39	Identities = 56/262 (21%), Positives = 121/262 (45%), Gaps = 28/262 (10%) gb AAF72344.1 AF192329_5 (AF192329) unknown [Enterococcus faecalis] Length = 287
SeqID 1146	SA-2100.1	Contig119 (3517-4416 p)	93	Identities = 263/299 (87%), Positives = 287/299 (95%) dbj BAA28714.1 (AB001562) glucose-1-phosphate uridylyltransferase [Streptococcus mutans] Length = 306
SeqID 1147	SA-2101.1	Contig119 (2464-3480 p)	66	Identities = 177/333 (53%), Positives = 241/333 (72%) sp P46919 GPDA_BACSU GLYCEROL-3-PHOSPHATE DEHYDROGENASE [NAD(P)+] (NAD(P)H-DEPENDENT DIHYDROXYACETONE-PHOSPHATE REDUCTASE) gb AAA86746.1 (U32164) NAD(P)H-dependent dihydroxyacetone-phosphate reductase [Bacillus subtilis] Length = 345
SeqID 1148	SA-2102.1	Contig119 (1965-2294 m)	54	Identities = 52/109 (47%), Positives = 77/109 (69%), Gaps = 1/109 (0%) pdb 1A6F Rnase P Protein From Bacillus Subtilis Length = 119
SeqID 1149	SA-2103.2	Contig119 (1137-1952 m)	71	Identities = 149/267 (55%), Positives = 197/267 (72%), Gaps = 3/267 (1%) gb AAK04227.1 AE006251_3 (AE006251) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 269
SeqID 1150	SA-2104.2	Contig133 (64785-66146 m)	39	Identities = 100/451 (22%), Positives = 180/451 (39%), Gaps = 81/451 (17%) emb CAB39034.1 (AL034559) hypothetical protein, PFC0940c [Plasmodium falciparum] Length = 806
SeqID 1151	SA-2105.1	Contig133 (66147-67238 m)	No Hits found	

SeqID 1152	SA-2106.1	Contig133 (67478-68422 m)	50	Identities = 100/317 (31%), Positives = 168/317 (52%), Gaps = 8/317 (2%) dbj BAB05608.1 (AP001513) unknown conserved protein [Bacillus halodurans] Length = 325
SeqID 1153	SA-2107.1	Contig133 (68486-68815 m)	No Hits found	
SeqID 1154	SA-2109.1	Contig133 (69774-70190 p)	No Hits found	
SeqID 1155	SA-211.1	Contig136 (54825-55058 m)	No Hits found	
SeqID 1156	SA-2110.1	Contig133 (70239-70367 p)	No Hits found	
SeqID 1157	SA-2112.1	Contig133 (70808-71098 p)	49	Identities = 35/91 (38%), Positives = 51/91 (55%) sp P34159 YHS1_CLOAB HYPOTHETICAL 11.0 KD PROTEIN IN HSP18 3 REGION (ORFA1) emb CAA46375.1 (X65276) ORFA1 [Clostridium acetobutylicum] Length = 96
SeqID 1158	SA-2113.1	Contig133 (71260-71571 p)	No Hits found	
SeqID 1159	SA-2115.1	Contig133 (71861-72589 p)	No Hits found	
SeqID 1160	SA-2116.1	Contig133 (72806-73078 p)	No Hits found	
SeqID 1161	SA-2117.1	Contig133 (73197-73502 m)	No Hits found	
SeqID 1162	SA-2118.2	Contig133 (73790-74143 m)	No Hits found	
SeqID 1163	SA-2119.2	Contig118 (5080-6015 p)	88	Identities = 260/311 (83%), Positives = 283/311 (90%) sp P95765 PPAC_STRGC PROBABLE MANGANESE-DEPENDENT INORGANIC PYROPHOSPHATASE (PYROPHOSPHATE PHOSPHO-HYDROLASE) (PPASE) gb AAB39104.1 (U57759) intragenetic coaggregation-relevant adhesin [Streptococcus gordonii] Length = 311
SeqID 1164	SA-212.1	Contig136 (52523-54868 m)	54	Identities = 261/793 (32%), Positives = 436/793 (54%), Gaps = 36/793 (4%) gb AAF72347.1 AF192329_8 (AF192329) TrsE-like protein [Enterococcus faecalis] Length = 799
SeqID 1165	SA-2120.1	Contig118 (4175-4963 p)	80	Identities = 185/260 (71%), Positives = 218/260 (83%) sp O68575 PFLA_STRMU PYRUVATE FORMATE-LYASE ACTIVATING ENZYME (PFL-ACTIVATING ENZYME) gb AAC05773.1 (AF051356) pyruvate-formate lyase activating enzyme [Streptococcus mutans] dbj BAA34998.1 (AB018417) PFL-activating enzyme [Streptococcus mutans] Length = 263

SeqID 1166	SA-2122.1	- Contig118 (2773-4107 p)	90	Identities = 348/445 (78%), Positives = 404/445 (90%), Gaps = 1/445 (0%) gb AAC05772.1 (AF051356) putative hemolysin [Streptococcus mutans] Length = 445
SeqID 1167	SA-2123.1	Contig118 (2027-2593 p)	62	Identities = 80/184 (43%), Positives = 119/184 (64%), Gaps = 3/184 (1%) pir G81942 hypothetical protein NMA0960 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB84230.1 (AL162754) hypothetical protein NMA0960 [Neisseria meningitidis] Length = 188
SeqID 1168	SA-2124.1	Contig118 (1099-2034 p)	69	Identities = 161/302 (53%), Positives = 220/302 (72%), Gaps = 4/302 (1%) pir D69999 conserved hypothetical protein ytcA - Bacillus subtilis gb AAC00380.1 (AF008220) YtcA [Bacillus subtilis] emb CAB15026.1 (Z99119) similar to hypothetical proteins [Bacillus subtilis] Length = 322
SeqID 1169	SA-2125.1	Contig118 (365-1006 p)	42	Identities = 62/159 (38%), Positives = 93/159 (57%), Gaps = 3/159 (1%) pir S32217 hypothetical protein 2 - Bacillus megaterium emb CAA79986.1 (Z21972) ORF2 [Bacillus megaterium] Length = 216
SeqID 1170	SA-2126.1	Contig118 (1-384 p)	48	Identities = 33/113 (29%), Positives = 62/113 (54%) sp P50726 YPAA_BACSU HYPOTHETICAL 20.5 KD PROTEIN IN SERA-FER INTERGENIC REGION pir E69932 hypothetical protein ypaA - Bacillus subtilis gb AAC83944.1 (L47648) putative [Bacillus subtilis] emb CAB14237.1 (Z99116) ypaA [Bacillus subtilis] Length = 190
SeqID 1171	SA-2128.2	Contig92 (6606-6896 p)	51	Identities = 44/74 (59%), Positives = 53/74 (71%) pir T44088 probable transposase [imported] - Staphylococcus aureus Length = 74

SeqID 1172	SA-2129.1	Contig92 (7096-7434 p)	60	Identities = 64/135 (47%), Positives = 86/135 (63%), Gaps = 1/135 (0%) ref NP_052792.1 pXO1-96 [Bacillus anthracis] pir H59102 hypothetical protein pXO1-96 - Bacillus anthracis virulence plasmid pXO1 gb AAA74027.1 (U30714) ORFB [Bacillus anthracis] gb AAA74029.1 (U30715) ORFB [Bacillus anthracis] gb AAD32400.1 AAD32400 (AF065404) pXO1-96 [Bacillus anthracis] Length = 274
SeqID 1173	SA-2130.1	Contig92 (7421-7738 p)	62	Identities = 45/98 (45%), Positives = 66/98 (66%) ref NP_052792.1 pXO1-96 [Bacillus anthracis] pir H59102 hypothetical protein pXO1-96 - Bacillus anthracis virulence plasmid pXO1 gb AAA74027.1 (U30714) ORFB [Bacillus anthracis] gb AAA74029.1 (U30715) ORFB [Bacillus anthracis] gb AAD32400.1 AAD32400 (AF065404) pXO1-96 [Bacillus anthracis] Length = 274
SeqID 1174	SA-2131.1	Contig92 (7784-8017 p)	74	Identities = 69/72 (95%), Positives = 69/72 (95%) pir T09011 probable transposase TnpA - Streptococcus pyogenes (fragment) gb AAB92607.1 (AF026542) TnpA [Streptococcus pyogenes] Length = 364
SeqID 1175	SA-2132.1	Contig92 (8216-10684 m)	98	Identities = 809/822 (98%), Positives = 816/822 (98%) pir T46758 hypothetical 92.4K protein - Streptococcus agalactiae gb AAD13797.1 (AF062533) unknown [Streptococcus agalactiae] Length = 822
SeqID 1176	SA-2133.1	Contig92 (10697-11617 m)	99	Identities = 303/306 (99%), Positives = 304/306 (99%) pir T46757 lipoprotein lmb [validated] - Streptococcus agalactiae gb AAD13796.1 (AF062533) Lmb [Streptococcus agalactiae] Length = 306
SeqID 1177	SA-2134.2	Contig139 (98649-99302 p)	79	Identities = 131/218 (60%), Positives = 176/218 (80%), Gaps = 1/218 (0%) gb AAD25108.1 AF140356_1 (AF140356) VncR [Streptococcus pneumoniae] emb CAB54582.1 (AJ006399) response regulator [Streptococcus pneumoniae] Length = 218

SeqID 1178	SA-2136.1	Contig139 (99299-100618 p)	63	Identities = 178/435 (40%), Positives = 281/435 (63%), Gaps = 1/435 (0%) gb AAD25109.1 AF140356_2 (AF140356) VncS [Streptococcus pneumoniae] emb CAB54583.1 (AJ006399) histidine kinase [Streptococcus pneumoniae] Length = 442
SeqID 1179	SA-2137.1	Contig139 (100670-101284 m)	90	Identities = 197/209 (94%), Positives = 200/209 (95%) pir T09011 probable transposase TnpA - Streptococcus pyogenes (fragment) gb AAB92607.1 (AF026542) TnpA [Streptococcus pyogenes] Length = 364
SeqID 1180	SA-2138.1	Contig139 (101495-101695 p)	No Hits found	
SeqID 1181	SA-2139.1	Contig139 (101737-101925 p)	No Hits found	
SeqID 1182	SA-214.1	Contig136 (49726-52500 m)	6	Identities = 45/117 (38%), Positives = 60/117 (50%), Gaps = 10/117 (8%) gb AAC61959.1 (AF051917) putative membrane protein TraG [Staphylococcus aureus] prf 2004267H traG protein [Staphylococcus sp.] Length = 358
SeqID 1183	SA-2140.1	Contig139 (102329-103555 p)	56	Identities = 142/392 (36%), Positives = 237/392 (60%), Gaps = 23/392 (5%) sp P39604 YWCF_BACSU HYPOTHETICAL 43.3 KD PROTEIN IN QOXD-VPR INTERGENIC REGION pir S39697 cell-division protein homolog ywcF - Bacillus subtilis emb CAA51598.1 (X73124) ipa-42d [Bacillus subtilis] emb CAB15838.1 (Z99123) alternate gene name: ipa-42d~similar to cell-division protein [Bacillus subtilis] Length = 393
SeqID 1184	SA-2141.1	Contig139 (103794-104234 p)	54	Identities = 82/142 (57%), Positives = 105/142 (73%) pir A57362 gyrb protein - Streptococcus pneumoniae (fragment) emb CAA58770.1 (X83917) orfgyrb [Streptococcus pneumoniae] emb CAA91552.1 (Z67740) unidentified [Streptococcus pneumoniae] Length = 144
SeqID 1185	SA-2143.2	Contig139 (104235-106187 p)	94	Identities = 574/650 (88%), Positives = 618/650 (94%), Gaps = 2/650 (0%) emb CAA91553.1 (Z67740) DNA gyrase [Streptococcus pneumoniae] Length = 648

SeqID 1186	SA-2145.1	Contig114 (17361-18035 m)	64	Identities = 98/223 (43%), Positives = 145/223 (64%), Gaps = 5/223 (2%) dbj BAB04091.1 (AP001508) two-component response regulator [Bacillus halodurans] Length = 222
SeqID 1187	SA-2146.1	Contig114 (18391-18525 m)	69	Identities = 33/44 (75%), Positives = 39/44 (88%) sp P23376 RL34_BACST 50S RIBOSOMAL PROTEIN L34 pir C48396 ribosomal protein L34 - Bacillus stearothermophilus gb AAB20570.1 BstL34=50S ribosomal subunit protein [Bacillus stearothermophilus, Peptide, 44 aa] gb AAB21085.1 ribosomal protein L34 [Bacillus stearothermophilus, Peptide, 44 aa] prf 1718186C ribosomal protein L34 [Bacillus stearothermophilus] Length = 44
SeqID 1188	SA-2147.1	Contig114 (18709-20064 m)	66	Identities = 211/459 (45%), Positives = 308/459 (66%), Gaps = 7/459 (1%) dbj BAB07666.1 (AP001520) unknown conserved protein [Bacillus halodurans] Length = 460
SeqID 1189	SA-2148.1	Contig114 (20311-22038 m)	73	Identities = 343/568 (60%), Positives = 426/568 (74%), Gaps = 2/568 (0%) gb AAF37879.1 AF234619_2 (AF234619) OpuABC [Lactococcus lactis] Length = 573
SeqID 1190	SA-2149.1	Contig114 (22057-23280 m)	81	Identities = 274/402 (68%), Positives = 337/402 (83%) gb AAF37878.1 AF234619_1 (AF234619) OpuAA [Lactococcus lactis] Length = 408
SeqID 1191	SA-2156.2	Contig98 (4979-5518 m)	62	Identities = 90/175 (51%), Positives = 118/175 (67%), Gaps = 2/175 (1%) sp P36264 NUSG_STACA TRANSCRIPTION ANTITERMINATION PROTEIN NUSG pir S38870 transcription antitermination factor nusG - Staphylococcus carnosus emb CAA53738.1 (X76134) nusG [Staphylococcus carnosus] Length = 182
SeqID 1192	SA-2157.1	Contig98 (3727-4932 p)	38	Identities = 98/259 (37%), Positives = 155/259 (59%), Gaps = 10/259 (3%) gb AAF28363.1 AF224467_2 (AF224467) putative glycosyl transferase [Haemophilus ducreyi] Length = 269

SeqID 1193	SA-2158.1	Contig98 (2466-3662 p)	37	Identities = 88/259 (33%), Positives = 156/259 (59%), Gaps = 11/259 (4%) gb AAF28363.1 AF224467_2 (AF224467) putative glycosyl transferase [Haemophilus ducreyi] Length = 269
SeqID 1194	SA-2159.1	Contig98 (1416-2225 m)	43	Identities = 75/260 (28%), Positives = 123/260 (46%), Gaps = 22/260 (8%) sp P39407 YJU_ECOLI_HYPOTHETICAL 39.8 KD PROTEIN IN OSMY-DEOC INTERGENIC REGION (O357) pir S56601 hypothetical 39.8K protein (osmY-deoC intergenic region) - Escherichia coli gb AAA97273.1 (U14003) ORF_o357 [Escherichia coli] gb AAC77330.1 (AE000508) orf, hypothetical protein [Escherichia coli K12] Length = 357
SeqID 1195	SA-216.1	Contig136 (48774-49622 p)	45	Identities = 84/265 (31%), Positives = 133/265 (49%), Gaps = 14/265 (5%) gb AAB52383.1 (U36837) AbiEii [Lactococcus lactis] Length = 298
SeqID 1196	SA-2160.1	Contig98 (126-1373 p)	52	Identities = 139/391 (35%), Positives = 221/391 (55%), Gaps = 4/391 (1%) sp P71369 YB04_HAEIN_HYPOTHETICAL METABOLITE TRANSPORT PROTEIN HI1104 pir C64167 hypothetical protein HI1104 - Haemophilus influenzae (strain Rd KW20) gb AAC22759.1 (U32790) transporter protein [Haemophilus influenzae Rd] Length = 407
SeqID 1197	SA-2161.1	Contig81 (8-1291 m)	11	Identities = 35/101 (34%), Positives = 52/101 (50%), Gaps = 1/101 (0%) pdb 1BU C Chain C, Structure Of The Ternary Microplasma-Staphylokinase-Microplasma Complex: A Proteinase-Cofactor-Substrate Complex In Action Length = 128
SeqID 1198	SA-2162.1	Contig81 (1758-2456 p)	No Hits found	

SeqID 1199	SA-2163.1	Contig81 (2627-3571 p)	46	Identities = 102/331 (30%), Positives = 172/331 (51%), Gaps = 26/331 (7%) sp P33019 YEIH_ECOLI_HYPOTHETICAL 36.9 KD PROTEIN IN LYSP-NFO INTERGENIC REGION pir E64984 hypothetical 36.9 kD protein in lysP-nfo intergenic region - Escherichia coli (strain K-12) gb AAA60511.1 (U00007) yeiH [Escherichia coli] gb AAC75219.1 (AE000305) orf, hypothetical protein [Escherichia coli K12] prf 2014253BD yeiH gene [Escherichia coli] Length = 349
SeqID 1200	SA-2165.1	Contig81 (3645-5021 p)	63	Identities = 229/455 (50%), Positives = 305/455 (66%), Gaps = 5/455 (1%) gb AAD19405.1 (AF102543) succinic semialdehyde dehydrogenase [Zymomonas mobilis] Length = 458
SeqID 1201	SA-2166.1	Contig81 (5164-5709 p)	68	Identities = 93/178 (52%), Positives = 127/178 (71%) sp P54417 OPUD_BACSU GLYCINE BETAIN TRANSPORTER OPUD pir G69670 glycine betaine transporter opuD - Bacillus subtilis gb AAC44368.1 (U50082) glycine betaine transporter OpuD [Bacillus subtilis] gb AAC00408.1 (AF008220) putative transporter [Bacillus subtilis] emb CAB14985.1 (Z99119) glycine betaine transporter [Bacillus subtilis] Length = 512
SeqID 1202	SA-2167.1	Contig81 (5710-6711 p)	70	Identities = 183/324 (56%), Positives = 236/324 (72%), Gaps = 1/324 (0%) pir T48645 glycine betaine transport protein betL [validated] - Listeria monocytogenes gb AAD30266.1 AF102174_1 (AF102174) glycine betaine transporter BetL [Listeria monocytogenes] Length = 507

SeqID 1203	SA-2168.1	Contig81 (6734-7219 m)	58	Identities = 68/152 (44%), Positives = 94/152 (61%), Gaps = 3/152 (1%) pir A70081 conserved hypothetical protein yxkH - Bacillus subtilis dbjBAA11724.1 (D83026) homologous to SwissProt:YADE_EC01; hypothetical [Bacillus subtilis] emb CAB15906.1 (Z99123) similar to hypothetical proteins [Bacillus subtilis] Length = 279
SeqID 1204	SA-2169.1	Contig81 (7060-7620 m)	29	Identities = 38/122 (31%), Positives = 56/122 (45%), Gaps = 17/122 (13%) pir A70081 conserved hypothetical protein yxkH - Bacillus subtilis dbjBAA11724.1 (D83026) homologous to SwissProt:YADE_EC01; hypothetical [Bacillus subtilis] emb CAB15906.1 (Z99123) similar to hypothetical proteins [Bacillus subtilis] Length = 279
SeqID 1205	SA-217.1	Contig136 (48187-48777 p)	44	Identities = 53/219 (24%), Positives = 93/219 (42%), Gaps = 30/219 (13%) gb AAB52382.1 (U36837) AbiEi [Lactococcus lactis] Length = 287
SeqID 1206	SA-2170.1	Contig81 (7797-8006 p)	58	Identities = 35/69 (50%), Positives = 50/69 (71%) spiP52985 DHOM_LACLA HOMOSERINE DEHYDROGENASE (HDH) pir JC6049 homoserine dehydrogenase (EC 1.1.1.3) - Lactococcus lactis emb CAA65713.1 (X96988) hom [Lactococcus lactis] Length = 428
SeqID 1207	SA-2172.1	Contig80 (5804-7288 p)	61	Identities = 220/470 (46%), Positives = 311/470 (65%), Gaps = 16/470 (3%) pir A82294 probable carbon starvation protein A VC0687 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAAF93852.1 (AE004154) carbon starvation protein A, putative [Vibrio cholerae] Length = 494
SeqID 1208	SA-2173.1	Contig80 (4914-5648 p)	59	Identities = 93/245 (37%), Positives = 150/245 (60%), Gaps = 3/245 (1%) gb AAB48183.1 (L42945) lytR [Staphylococcus aureus] Length = 246
SeqID 1209	SA-2174.1	Contig80 (3163-4902 p)	67	Identities = 265/582 (45%), Positives = 394/582 (67%), Gaps = 2/582 (0%) gb AAB48182.1 (L42945) lytS [Staphylococcus aureus] Length = 584
SeqID 1210	SA-2175.1	Contig80 (2618-2782 p)	No Hits found	

SeqID 1211	SA-2177.2	Contig80 (2239-2418 p)	No Hits found	
SeqID 1212	SA-2178.1	Contig80 (1241-1618 p)	No Hits found	
SeqID 1213	SA-2180.2	Contig80 (529-852 p)	No Hits found	
SeqID 1214	SA-2182.2	Contig80 (2-181 p)	No Hits found	
SeqID 1215	SA-2184.1	Contig129 (22347-23933 m)	66	Identities = 283/489 (57%), Positives = 361/489 (72%), Gaps = 14/489 (2%) gb AAD20136.1 (AF091502) autoaggregation-mediating protein [Lactobacillus reuteri] Length = 497
SeqID 1216	SA-2185.1	Contig129 (24168-24998 m)	46	Identities = 89/237 (37%), Positives = 132/237 (55%), Gaps = 3/237 (1%) pir C81348 probable periplasmic protein Cj0771c [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73036.1 (AL139076) putative periplasmic protein [Campylobacter jejuni] Length = 256
SeqID 1217	SA-2186.1	Contig129 (25014-25676 m)	56	Identities = 80/206 (38%), Positives = 127/206 (60%), Gaps = 4/206 (1%) pir D82957 probable permease of ABC transporter PA5504 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG08889.1 AE004963_2 (AE004963) probable permease of ABC transporter [Pseudomonas aeruginosa] Length = 225
SeqID 1218	SA-2187.1	Contig129 (25669-26403 m)	56	Identities = 84/221 (38%), Positives = 138/221 (62%), Gaps = 5/221 (2%) sp P44785 ABC_HAEIN ATP-BINDING PROTEIN ABC pir C64082 ATP-binding protein homolog HI0621 - Haemophilus influenzae (strain Rd KW20) gb AAC22280.1 (U32744) ABC transporter, ATP-binding protein [Haemophilus influenzae Rd] Length = 345
SeqID 1219	SA-2188.1	Contig129 (26524-26904 m)	69	Identities = 74/125 (59%), Positives = 92/125 (73%) pir J69854 hypothetical protein yjqa - Bacillus subtilis emb CAB13104.1 (Z99110) yjqa [Bacillus subtilis] gb AAB87515.1 (AF034138) unknown [Bacillus subtilis] Length = 125

SeqID 1220	SA-2190.1	Contig129 (26990-28534 m)	75	Identities = 311/518 (60%), Positives = 393/518 (75%), Gaps = 9/518 (1%) sp O86490 RF3_STAAU PEPTIDE CHAIN RELEASE FACTOR 3 (RF-3) emb CAA74739.1 (Y14370) peptide chain release factor 3 [Staphylococcus aureus] Length = 521
SeqID 1221	SA-2192.2	Contig129 (28711-30249 m)	26	Identities = 92/358 (25%), Positives = 142/358 (38%), Gaps = 55/358 (15%) gb AAG54632.1 AE005207_7 (AE005207) putative adhesin [Escherichia coli O157:H7] Length = 1417
SeqID 1222	SA-2195.2	Contig109 (153-1748 m)	56	Identities = 243/373 (65%), Positives = 302/373 (80%), Gaps = 1/373 (0%) sp P34001 YWAP_STRMU HYPOTHETICAL PROTEIN IN WAPA 3 REGION pir S06993 hypothetical protein (wapA 3 region) - Streptococcus mutans (fragment) gb AAA88609.1 (M37842) unknown protein [Streptococcus mutans] Length = 373
SeqID 1223	SA-2196.1	Contig109 (1867-3537 m)	87	Identities = 432/556 (77%), Positives = 492/556 (87%) sp Q59925 FTHS_STRMU FORMATE--TETRAHYDROFOLATE LIGASE (FORMYL-TETRAHYDROFOLATE SYNTHETASE) (FHS) (FTHFS) gb AAB49329.1 (U39612) formyl-tetrahydrofolate synthetase [Streptococcus mutans] Length = 556
SeqID 1224	SA-2197.1	Contig109 (3626-4645 m)	60	Identities = 131/331 (39%), Positives = 207/331 (61%), Gaps = 5/331 (1%) pir G69830 lipoate-protein ligase homolog yhfJ - Bacillus subtilis emb CAA74531.1 (Y14083) hypothetical protein [Bacillus subtilis] emb CAB12865.1 (Z99109) similar to lipoate-protein ligase [Bacillus subtilis] Length = 331
SeqID 1225	SA-2198.2	Contig109 (4672-5550 m)	No Hits found	
SeqID 1226	SA-22.1	Contig137 (22578-24410 p)	52	Identities = 228/571 (39%), Positives = 347/571 (59%), Gaps = 24/571 (4%) pir E81869 probable ATP-dependent proteinase ATP-binding protein NMA1045 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB84311.1 (AL162755) putative ATP-dependent protease ATP-binding protein [Neisseria meningitidis] Length = 759

SeqID 1227	SA-220.1	Contig136 (43070-47860 m)	36	Identities = 420/1385 (30%), Positives = 606/1385 (43%), Gaps = 308/1385 (22%) gb AAC44100.1 (U40026) SspB precursor [Streptococcus gordonii] Length = 1500
SeqID 1228	SA-2200.2	Contig78 (4809-6092 p)	92	Identities = 365/427 (85%), Positives = 404/427 (94%) sp O85730 TIG_STRPY TRIGGER FACTOR (TF) gb AAC82391.1 (AF073922) RopA [Streptococcus pyogenes] Length = 427
SeqID 1229	SA-2201.1	Contig78 (3810-4652 m)	52	Identities = 105/261 (40%), Positives = 150/261 (57%), Gaps = 2/261 (0%) dbj BAB06385.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 276
SeqID 1230	SA-2202.1	Contig78 (3204-3773 p)	58	Identities = 91/176 (51%), Positives = 115/176 (64%) sp P39157 YWLG_BACSU HYPOTHETICAL 19.4 KD PROTEIN IN SPOIR-GLYC INTERGENIC REGION pir I40482 hypothetical protein ywlg - Bacillus subtilis emb CAA86109.1 (Z38002) Unknown [Bacillus subtilis] emb CAB15708.1 (Z99122) alternate gene name: ipc-33d [Bacillus subtilis] prf 2108403H ipc-33d gene [Bacillus subtilis] Length = 180
SeqID 1231	SA-2203.1	Contig78 (2743-3207 p)	49	Identities = 46/148 (31%), Positives = 78/148 (52%), Gaps = 9/148 (6%) pir G75153 hypothetical protein PAB2090 - Pyrococcus abyssi (strain Orsay) emb CAB49310.1 (AJ248284) hypothetical protein [Pyrococcus abyssi] Length = 199
SeqID 1232	SA-2204.1	Contig78 (1975-2733 p)	58	Identities = 95/253 (37%), Positives = 150/253 (58%), Gaps = 13/253 (5%) sp P39610 THID_BACSU PHOSPHOMETHYLPYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE) pir S39707 phosphomethylpyrimidine kinase thiD - Bacillus subtilis emb CAA51608.1 (X73124) ipa-52r [Bacillus subtilis] emb CAB15828.1 (Z99123) phosphomethylpyrimidine kinase [Bacillus subtilis] Length = 271

SeqID 1233	SA-2205.1	Contig78 (1236-2012 p)	56	Identities = 105/240 (43%), Positives = 147/240 (60%), Gaps = 2/240 (0%) sp Q9Z9J0 TRUA_BACHD TRNA PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE I) (PSEUDOURIDINE SYNTHASE I) (URACIL HYDROLYASE) pir T44415 pseudouridylylate synthase I truA [imported] - Bacillus halodurans dbj BAA75303.1 (AB017508) truA homologue (identity of 62 to B. subtilis%) [Bacillus halodurans] dbj BAB03886.1 (AP001507) tRNA pseudouridine synthase A (pseudouridylylate synthase I) [Bacillus halodurans] Length = 263
SeqID 1234	SA-2206.1	Contig78 (2-1165 p)	62	Identities = 173/347 (49%), Positives = 241/347 (68%), Gaps = 3/347 (0%) gb AAD24445.1 AF118389_2 (AF118389) unknown [Streptococcus suis] Length = 419
SeqID 1235	SA-2207.2	Contig139 (46557-47384 p)	73	Identities = 162/270 (60%), Positives = 202/270 (74%), Gaps = 3/270 (1%) dbj BAB06497.1 (AP001516) hemolysin-like protein [Bacillus halodurans] Length = 272
SeqID 1236	SA-2208.2	Contig139 (47371-47844 p)	50	Identities = 49/153 (32%), Positives = 84/153 (54%), Gaps = 4/153 (2%) emb CAA09426.1 (AJ010954) arginine repressor [Bacillus stearothermophilus] Length = 149
SeqID 1237	SA-2210.1	Contig139 (47856-49514 p)	63	Identities = 245/567 (43%), Positives = 366/567 (64%), Gaps = 18/567 (3%) sp P17894 REC_N_BACSU DNA REPAIR PROTEIN REC_N (RECOMBINATION PROTEIN N) pir B35128 DNA repair and genetic recombination protein recN - Bacillus subtilis gb AA22691.1 (M30297) recombination protein (ttg start codon) [Bacillus subtilis] dbj BAA12579.1 (D84432) RecN [Bacillus subtilis] emb CAB14355.1 (Z99116) recN [Bacillus subtilis] Length = 576
SeqID 1238	SA-2212.1	Contig139 (49627-50463 p)	54	Identities = 93/277 (33%), Positives = 152/277 (54%), Gaps = 4/277 (1%) dbj BAB07346.1 (AP001519) unknown conserved protein [Bacillus halodurans] Length = 283

SeqID 1239	SA-2213.1	Contig139 (50531-51295 p)	69	Identities = 149/274 (54%), Positives = 208/274 (75%), Gaps = 9/274 (3%) emb CAA72096.1 (Y11213) hypothetical protein [Streptococcus thermophilus] Length = 280
SeqID 1240	SA-2214.2	Contig139 (51270-51872 p)	57	Identities = 75/185 (40%), Positives = 116/185 (62%), Gaps = 3/185 (1%) emb CAA72097.1 (Y11213) hypothetical protein [Streptococcus thermophilus] Length = 189
SeqID 1241	SA-2215.1	Contig77 (5615-5863 m)	55	Identities = 31/84 (36%), Positives = 51/84 (59%) pir B69770 conserved hypothetical protein ydaS - Bacillus subtilis dbj BAA19274.1 (AB001488) FUNCTION UNKNOWN. [Bacillus subtilis] emb CAB12244.1 (Z99106) similar to hypothetical proteins from B. subtilis [Bacillus subtilis] Length = 85
SeqID 1242	SA-2216.1	Contig77 (5019-5570 m)	No Hits found	
SeqID 1243	SA-2217.1	Contig77 (4815-5009 m)	66	Identities = 27/61 (44%), Positives = 45/61 (73%) gb AAA86382.1 (U23376) putative 6-kDa protein [Lactococcus lactis] Length = 62
SeqID 1244	SA-2219.1	Contig77 (4217-4759 m)	64	Identities = 95/157 (60%), Positives = 121/157 (76%) gb AAA86383.1 (U23376) putative 20-kDa protein [Lactococcus lactis] Length = 183
SeqID 1245	SA-222.1	Contig136 (42878-43069 m)	45	Identities = 20/43 (46%), Positives = 29/43 (66%), Gaps = 1/43 (2%) gb AAG19662.1 (AF005054) calcium-binding protein homology; Cbp [Halobacterium sp. NRC-1] Length = 385
SeqID 1246	SA-2220.1	Contig77 (3961-4158 m)	53	Identities = 24/61 (39%), Positives = 37/61 (60%), Gaps = 1/61 (1%) gb AAB96651.1 (AF034574) putative cruciform DNA binding protein [Glomus versiforme] Length = 99
SeqID 1247	SA-2221.1	Contig77 (3388-3939 m)	58	Identities = 83/153 (54%), Positives = 110/153 (71%) gb AAA86383.1 (U23376) putative 20-kDa protein [Lactococcus lactis] Length = 183

SeqID 1248	SA-2222.1	Contig77 (2583-3224 m)	15	<p>Identities = 22/48 (45%), Positives = 35/48 (72%) ref NP_070072.1 A. fulgidus predicted coding region AF1244 [Archaeoglobus fulgidus] pir C69405 hypothetical protein AF1244 - Archaeoglobus fulgidus gb AAB90005.1 (AE001018) A. fulgidus predicted coding region AF1244 [Archaeoglobus fulgidus] Length = 161</p>
SeqID 1249	SA-2223.1	Contig77 (1814-2578 m)	46	<p>Identities = 67/262 (25%), Positives = 121/262 (45%), Gaps = 21/262 (8%) pir B72352 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD35735.1 AE001738_15 (AE001738) conserved hypothetical protein [Thermotoga maritima] Length = 268</p>
SeqID 1250	SA-2224.1	Contig77 (1155-1814 m)	52	<p>Identities = 67/218 (30%), Positives = 120/218 (54%), Gaps = 5/218 (2%) emb CAB40581.1 (AJ010128) DNA alkylation repair enzyme [Bacillus cereus] Length = 237</p>
SeqID 1251	SA-2225.1	Contig77 (588-1082 m)	85	<p>Identities = 119/163 (73%), Positives = 145/163 (88%) sp P31308 TPX_STRSA PROBABLE THIOL PEROXIDASE pir B43583 thioredoxin peroxidase (EC 1.11.1.-) - Streptococcus sanguis gb AAC98427.1 (M63481) 20-kDa protein [Streptococcus sanguinis] Length = 163</p>
SeqID 1252	SA-2226.1	Contig77 (3-512 p)	No Hits found	
SeqID 1253	SA-2227.2	Contig133 (18629-19093 m)	46	<p>Identities = 39/135 (28%), Positives = 76/135 (55%), Gaps = 9/135 (6%) gb AAG09977.1 AF248038_6 (AF248038) Gata [Streptococcus agalactiae] Length = 149</p>
SeqID 1254	SA-2228.1	Contig133 (18322-18627 m)	No Hits found	
SeqID 1255	SA-2231.1	Contig133 (16834-18282 m)	40	<p>Identities = 112/408 (27%), Positives = 197/408 (47%), Gaps = 21/408 (5%) sp P39365 SGCC_ECOLI PUTATIVE PHOSPHOTRANSFERASE ENZYME II, C COMPONENT SGCC pir S56529 probable phosphotransferase enzyme II - Escherichia coli gb AAA97200.1 (U14003) ORF_f437 [Escherichia coli] gb AAC77260.1 (AE000501) putative PTS system enzyme IIC component [Escherichia coli K12] Length = 437</p>

SeqID 1256	SA-2232.2	Contig133 (14146-16596 m)	45	Identities = 294/504 (58%), Positives = 380/504 (75%), Gaps = 10/504 (1%) sp Q59959 NANA_STRPN SIALIDASE A PRECURSOR (NEURAMINIDASE A) pir T30287 exo-alpha-sialidase (EC 3.2.1.18) - Streptococcus pneumoniae emb CAA51473.1 (X72967) neuraminidase [Streptococcus pneumoniae] Length = 1035
SeqID 1257	SA-2233.2	Contig101 (6129-7064 p)	25	Identities = 57/146 (39%), Positives = 81/146 (55%), Gaps = 8/146 (5%) gb AAF18951.1 AF155805_5 (AF155805) Cps9H [Streptococcus suis] Length = 143
SeqID 1258	SA-2235.1	Contig101 (5072-6127 p)	47	Identities = 108/327 (33%), Positives = 170/327 (51%), Gaps = 22/327 (6%) emb CAC14890.1 (AJ295156) d-TDP-glucose dehydratase [Phragmites australis] Length = 350
SeqID 1259	SA-2236.1	Contig101 (4347-5069 p)	52	Identities = 77/231 (33%), Positives = 131/231 (56%), Gaps = 6/231 (2%) pir S66119 conserved hypothetical protein yacM - Bacillus subtilis dbj BAA05324.1 (D26185) unknown [Bacillus subtilis] emb CAB11866.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 232
SeqID 1260	SA-2237.1	Contig101 (3523-4347 p)	50	Identities = 85/274 (31%), Positives = 141/274 (51%), Gaps = 16/274 (5%) gb AAD37093.1 AF106539_2 (AF106539) LicD1 [Streptococcus pneumoniae] Length = 267
SeqID 1261	SA-2238.1	Contig101 (1765-3498 p)	No Hits found	
SeqID 1262	SA-2239.1	Contig101 (1419-1772 p)	46	Identities = 27/109 (24%), Positives = 55/109 (49%) dbj BAA19645.1 (AB002668) unnamed protein product [Actinobacillus actinomycetemcomitans] Length = 126
SeqID 1263	SA-224.1	Contig136 (42343-42894 m)	No Hits found	
SeqID 1264	SA-2240.1	Contig101 (691-1422 p)	59	Identities = 118/240 (49%), Positives = 152/240 (63%), Gaps = 1/240 (0%) gb AAC35924.1 (AF071085) putative glycosyl transferase [Enterococcus faecalis] Length = 241

SeqID 1265	SA-2241.2	Contig101 (3-689 p)	75	Identities = 123/231 (53%), Positives = 174/231 (75%), Gaps = 7/231 (3%) pir T00087 rhamnosyltransferase - Streptococcus mutans dbj BAA32090.1 (AB010970) rhamnosyltransferase [Streptococcus mutans] Length = 311
SeqID 1266	SA-2242.2	Contig115 (41-406 p)	75	Identities = 77/118 (65%), Positives = 102/118 (86%) gb AAK04289.1 AE006256_11 (AE006256) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 122
SeqID 1267	SA-2244.2	Contig115 (406-2070 p)	78	Identities = 354/539 (65%), Positives = 438/539 (80%), Gaps = 8/539 (1%) gb AAK04288.1 AE006256_10 (AE006256) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 535
SeqID 1268	SA-2245.1	Contig115 (2254-3102 p)	56	Identities = 127/275 (46%), Positives = 174/275 (63%), Gaps = 1/275 (0%) gb AAF68390.1 AF236374_1 (AF236374) hypersensitive-induced response protein [Zea mays] Length = 284
SeqID 1269	SA-2246.1	Contig115 (4134-4664 p)	31	Identities = 39/110 (35%), Positives = 55/110 (49%), Gaps = 3/110 (2%) pir G72536 hypothetical protein APE1580 - Aeropyrum pernix (strain K1) dbj BAA80580.1 (AP000062) 114aa long hypothetical protein [Aeropyrum pernix] Length = 114
SeqID 1270	SA-2247.1	Contig115 (4184-4924 m)	71	Identities = 132/241 (54%), Positives = 178/241 (73%), Gaps = 1/241 (0%) pir F81363 probable glutamine transport ATP-binding protein Cj0902 [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73160.1 (AL139076) putative glutamine transport ATP-binding protein [Campylobacter jejuni] Length = 242
SeqID 1271	SA-2248.2	Contig115 (4934-6484 m)	49	Identities = 147/534 (27%), Positives = 255/534 (47%), Gaps = 75/534 (14%) pir S77250 hypothetical protein - Synechocystis sp. (strain PCC 6803) dbj BAA17584.1 (D90907) glutamine-binding periplasmic protein [Synechocystis sp.] Length = 530
SeqID 1272	SA-225.1	Contig136 (41699-42292 m)	No Hits found	

SeqID 1273	SA-2251.1	Contig104 (9163-9489 p)	47	<p>Identities = 35/95 (36%), Positives = 56/95 (58%), Gaps = 3/95 (3%) sp P46339 YQGH_BACSU PROBABLE ABC TRANSPORTER PERMEASE PROTEIN YQGH pir B69956 phosphate ABC transporter (permease) homolog yqgH - Bacillus subtilis dbj BAA09582.1 (D58414) ORF72 [Bacillus subtilis] dbj BAA12511.1 (D84432) YqgH [Bacillus subtilis] emb CAB14428.1 (Z99116) alternate gene name: yzmC~similar to phosphate ABC transporter (permease) [Bacillus subtilis] Length = 309</p>
SeqID 1274	SA-2252.1	Contig104 (9422-9700 p)	45	<p>Identities = 35/54 (64%), Positives = 44/54 (80%) sp P46339 YQGH_BACSU PROBABLE ABC TRANSPORTER PERMEASE PROTEIN YQGH pir B69956 phosphate ABC transporter (permease) homolog yqgH - Bacillus subtilis dbj BAA09582.1 (D58414) ORF72 [Bacillus subtilis] dbj BAA12511.1 (D84432) YqgH [Bacillus subtilis] emb CAB14428.1 (Z99116) alternate gene name: yzmC~similar to phosphate ABC transporter (permease) [Bacillus subtilis] Length = 309</p>
SeqID 1275	SA-2253.1	Contig104 (9663-10082 p)	62	<p>Identities = 78/161 (48%), Positives = 113/161 (69%), Gaps = 1/161 (0%) sp P46339 YQGH_BACSU PROBABLE ABC TRANSPORTER PERMEASE PROTEIN YQGH pir B69956 phosphate ABC transporter (permease) homolog yqgH - Bacillus subtilis dbj BAA09582.1 (D58414) ORF72 [Bacillus subtilis] dbj BAA12511.1 (D84432) YqgH [Bacillus subtilis] emb CAB14428.1 (Z99116) alternate gene name: yzmC~similar to phosphate ABC transporter (permease) [Bacillus subtilis] Length = 309</p>

SeqID 1276	SA-2254.1	Contig104 (10072-10959 p)	74	<p>Identities = 157/294 (53%), Positives = 225/294 (76%) sp P46340 YQGI_BACSU PROBABLE ABC TRANSPORTER PERMEASE PROTEIN YQGI pir C69956 phosphate ABC transporter (permease) homolog yqgl - Bacillus subtilis dbj BAA09583.1 (D58414) ORF73 [Bacillus subtilis] dbj BAA12512.1 (D84432) Yqgl [Bacillus subtilis] emb CAB14427.1 (Z99116) alternate gene name: yzmD~similar to phosphate ABC transporter (permease) [Bacillus subtilis] Length = 294</p>
SeqID 1277	SA-2255.1	Contig104 (10971-11774 p)	74	<p>Identities = 154/247 (62%), Positives = 204/247 (82%) sp Q58418 PSTB_METJA PROBABLE PHOSPHATE TRANSPORT ATP-BINDING PROTEIN PSTB pir C64426 phosphate transport system ATP-binding protein - Methanococcus jannaschii gb AAB99016.1 (U67544) phosphate specific transport complex component (pstB) [Methanococcus jannaschii] Length = 252</p>
SeqID 1278	SA-2256.1	Contig104 (11786-12544 p)	74	<p>Identities = 148/248 (59%), Positives = 189/248 (75%) sp P46341 YQGJ_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YQGJ pir D69956 phosphate ABC transporter (ATP-binding pro) homolog yqgJ - Bacillus subtilis dbj BAA09584.1 (D58414) ORF74 [Bacillus subtilis] dbj BAA12513.1 (D84432) YqgJ [Bacillus subtilis] emb CAB14426.1 (Z99116) alternate gene name: yzmE~similar to phosphate ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 269</p>
SeqID 1279	SA-2258.2	Contig104 (12578-13231 p)	72	<p>Identities = 116/217 (53%), Positives = 167/217 (76%) gb AAK05813.1 AE006402_1 (AE006402) phosphate transport system regulator [Lactococcus lactis subsp. lactis] Length = 217</p>

SeqID 1280	SA-2259.1	Contig76 (2249-4879 p)	25	Identities = 107/531 (20%), Positives = 224/531 (42%), Gaps = 62/531 (11%) ref NP_070647.1 A. fulgidus predicted coding region AF1820 [Archaeoglobus fulgidus] pir C69477 hypothetical protein AF1820 - Archaeoglobus fulgidus gb AAB89436.1 (AE000977) A. fulgidus predicted coding region AF1820 [Archaeoglobus fulgidus] Length = 791
SeqID 1281	SA-2260.1	Contig76 (1536-2237 p)	70	Identities = 112/230 (48%), Positives = 167/230 (71%) ref NP_070646.1 ABC transporter, ATP-binding protein [Archaeoglobus fulgidus] pir B69477 ABC transporter, ATP-binding protein homolog - Archaeoglobus fulgidus gb AAB89431.1 (AE000977) ABC transporter, ATP-binding protein [Archaeoglobus fulgidus] Length = 231
SeqID 1282	SA-2261.1	Contig76 (158-1399 p)	76	Identities = 278/469 (59%), Positives = 355/469 (75%), Gaps = 10/469 (2%) sp Q9KA23 TOP1_BACHD DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE) dbj BAB06186.1 (AP001515) DNA topoisomerase I [Bacillus halodurans] Length = 690
SeqID 1283	SA-2263.1	Contig127 (39359-39559 p)	No Hits found	
SeqID 1284	SA-2264.2	Contig127 (39121-39972 m)	64	Identities = 126/284 (44%), Positives = 185/284 (64%) sp P37550 ISPE_BACSU 4-DIPHOSPHOCYTIDYL-2-C-METHYL-D-ERYTHRITOL KINASE (CMK) (4-CYTIDINE-5-DIPHOSPHO)-2-C-METHYL-D-ERYTHRITOL KINASE) pir S66075 conserved hypothetical protein yabH - Bacillus subtilis dbj BAA05281.1 (D26185) unknown [Bacillus subtilis] emb CAB11822.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 289
SeqID 1285	SA-2265.2	Contig127 (38592-39035 m)	73	Identities = 77/146 (52%), Positives = 117/146 (79%) pir T46753 repressor protein adcR [imported] - Streptococcus pneumoniae emb CAA96184.1 (Z71552) AdcR protein [Streptococcus pneumoniae] Length = 146

SeqID 1286	SA-2266.2	Contig127 (37879-38589 m)	84	Identities = 182/231 (78%), Positives = 206/231 (88%) pir T46754 AdcC-protein [imported] - Streptococcus pneumoniae emb CAA96186.1 (Z71552) AdcC protein [Streptococcus pneumoniae] Length = 234
SeqID 1287	SA-2267.1	Contig127 (37077-37889 m)	77	Identities = 197/263 (74%), Positives = 236/263 (88%) pir T46755 membrane protein adcB [imported] - Streptococcus pneumoniae emb CAA96187.1 (Z71552) AdcB protein [Streptococcus pneumoniae] Length = 268
SeqID 1288	SA-2268.2	Contig127 (35834-36865 p)	41	Identities = 116/216 (53%), Positives = 150/216 (68%), Gaps = 9/216 (4%) gb AAK04254.1 AE006253_5 (AE006253) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 263
SeqID 1289	SA-2269.1	Contig134 (81376-82521 m)	61	Identities = 169/374 (45%), Positives = 235/374 (62%), Gaps = 16/374 (4%) sp O06005 AAPA_BACSU AMINO ACID PERMEASE AAPA pir B69580 amino acid permease aapA - Bacillus subtilis emb CAA63459.1 (X92868) amino acid permease [Bacillus subtilis] emb CAB14651.1 (Z99117) amino acid permease [Bacillus subtilis] Length = 459
SeqID 1290	SA-2271.1	Contig134 (80825-81292 p)	86	Identities = 121/155 (78%), Positives = 139/155 (89%) gb AAC23745.1 (AF052209) VacB homolog [Streptococcus pneumoniae] Length = 441
SeqID 1291	SA-2272.2	Contig134 (78417-80822 p)	61	Identities = 350/815 (42%), Positives = 501/815 (60%), Gaps = 49/815 (6%) sp O32231 RNR_BACSU RIBONUCLEASE R (RNASE R) (VACB PROTEIN HOMOLOG) pir G70027 conserved hypothetical protein yvaJ - Bacillus subtilis emb CAB15366.1 (Z99121) similar to hypothetical proteins [Bacillus subtilis] Length = 779
SeqID 1292	SA-2274.1	Contig123 (1-786 m)	76	Identities = 169/260 (65%), Positives = 210/260 (80%), Gaps = 1/260 (0%) gb AAF86640.1 AF162694_1 (AF162694) ABC transporter [Enterococcus gallinarum] Length = 269

SeqID 1293	SA-2276.1	Contig123 (791-1687 m)	57	Identities = 114/285 (40%), Positives = 175/285 (61%), Gaps = 3/285 (1%) pir F83165 probable permease of ABC transporter PA3837 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG07224.1 AE004801_2 (AE004801) probable permease of ABC transporter [Pseudomonas aeruginosa] Length = 296
SeqID 1294	SA-2278.1	Contig123 (1703-2647 m)	50	Identities = 103/283 (36%), Positives = 165/283 (57%), Gaps = 1/283 (0%) pir E83165 hypothetical protein PA3836 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG07223.1 AE004801_1 (AE004801) hypothetical protein [Pseudomonas aeruginosa] Length = 325
SeqID 1295	SA-2279.1	Contig123 (3130-3939 m)	54	Identities = 89/261 (34%), Positives = 151/261 (57%), Gaps = 4/261 (1%) sp P44447 Y003_HAEIN PROTEIN HI0003 pir I64139 hypothetical protein HI0003 - Haemophilus influenzae (strain Rd KW20) gb AAC21682.1 (U32686) conserved hypothetical protein [Haemophilus influenzae Rd] Length = 262
SeqID 1296	SA-2280.1	Contig123 (4113-4973 p)	43	Identities = 86/275 (31%), Positives = 143/275 (51%), Gaps = 6/275 (2%) dbj BAB07024.1 (AP001518) unknown conserved protein [Bacillus halodurans] Length = 286
SeqID 1297	SA-2281.2	Contig123 (5016-5747 p)	58	Identities = 97/240 (40%), Positives = 151/240 (62%), Gaps = 1/240 (0%) gb AAD24446.1 AF118389_3 (AF118389) unknown [Streptococcus suis] Length = 244
SeqID 1298	SA-2282.1	Contig104 (20416-20709 p)	No Hits found	
SeqID 1299	SA-2283.1	Contig104 (20191-20379 p)	No Hits found	
SeqID 1300	SA-2285.1	Contig104 (18504-20069 p)	92	Identities = 437/522 (83%), Positives = 484/522 (92%), Gaps = 7/522 (1%) sp Q54431 SR54_STRMU SIGNAL RECOGNITION PARTICLE PROTEIN (FIFTY-FOUR HOMOLOG) gb AAB48050.1 AAB48050 (U88582) Ffh [Streptococcus mutans] Length = 516
SeqID 1301	SA-2287.1	Contig104 (18154-18486 p)	84	Identities = 95/110 (86%), Positives = 103/110 (93%) gb AAB48049.1 AAB48049 (U88582) YlxM [Streptococcus mutans] Length = 110

SeqID 1302	SA-2288.2	Contig104 (16752-18065 p)	69	Identities = 219/420 (52%), Positives = 306/420 (72%), Gaps = 4/420 (0%) sp Q54955 CIAH_STRPN SENSOR PROTEIN CIAH pir S49545 histidine kinase - Streptococcus pneumoniae emb CAA54466.1 (X77249) histidine kinase [Streptococcus pneumoniae] emb CAB54565.1 (AJ05926) histidine kinase [Streptococcus pneumoniae] Length = 444
SeqID 1303	SA-2289.2	Contig129 (21139-21903 p)	73	Identities = 152/245 (62%), Positives = 193/245 (78%), Gaps = 1/245 (0%) gb AAK06239.1 AE006442_6 (AE006442) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 250
SeqID 1304	SA-2291.1	Contig129 (20264-21004 m)	44	Identities = 66/200 (33%), Positives = 111/200 (55%), Gaps = 10/200 (5%) pir E69826 probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) yhdO - Bacillus subtilis emb CAA74499.1 (Y14082) hypothetical protein [Bacillus subtilis] emb CAB12793.1 (Z99109) similar to 1-acylglycerol-3-phosphate O-acyltransferase [Bacillus subtilis] Length = 199
SeqID 1305	SA-2292.1	Contig129 (19511-20164 m)	60	Identities = 96/217 (44%), Positives = 138/217 (63%), Gaps = 4/217 (1%) gb AAC23741.1 (AF052208) competence protein [Streptococcus pneumoniae] Length = 216
SeqID 1306	SA-2293.1	Contig129 (18655-19527 m)	60	Identities = 120/286 (41%), Positives = 180/286 (61%), Gaps = 1/286 (0%) gb AAC23742.1 (AF052208) competence protein [Streptococcus pneumoniae] Length = 753
SeqID 1307	SA-2294.1	Contig129 (17289-18554 m)	64	Identities = 205/419 (48%), Positives = 298/419 (70%), Gaps = 2/419 (0%) gb AAC23742.1 (AF052208) competence protein [Streptococcus pneumoniae] Length = 753
SeqID 1308	SA-2295.2	Contig129 (16354-17163 m)	61	Identities = 120/267 (44%), Positives = 177/267 (65%), Gaps = 6/267 (2%) gb AAK04342.1 AE006262_1 (AE006262) conserved hypothetical protein [Lactococcus lactis subsp. lactis] Length = 270

SeqID 1309	SA-2296.2	Contig101 (10807-11655 p)	57	Identities = 112/278 (40%), Positives = 163/278 (58%), Gaps = 2/278 (0%) emb CAB52237.1 (Z98171) EpsQ protein [Streptococcus thermophilus] Length = 279
SeqID 1310	SA-2297.1	Contig101 (11645-12784 p)	51	Identities = 108/395 (27%), Positives = 196/395 (49%), Gaps = 28/395 (7%) pir D64433 hypothetical protein MJ1069 - Methanococcus jannaschii gb AAB99071.1 (U67549) galactosyltransferase isolog [Methanococcus jannaschii] Length = 392
SeqID 1311	SA-2298.1	Contig101 (12836-14365 m)	54	Identities = 172/492 (34%), Positives = 289/492 (57%), Gaps = 2/492 (0%) gb AAC97147.1 (U49397) Nra [Streptococcus pyogenes] Length = 511
SeqID 1312	SA-23.1	Contig137 (20828-22576 p)	53	Identities = 218/608 (35%), Positives = 323/608 (52%), Gaps = 52/608 (8%) gb AAC38606.1 (AF007787) type I topoisomerase [Enterococcus faecalis] Length = 714
SeqID 1313	SA-230.1	Contig136 (35495-41695 m)	29	Identities = 373/1243 (30%), Positives = 614/1243 (49%), Gaps = 97/1243 (7%) ref NP_066674.1 similar to jhp0928 gene in Helicobacter pylori [Agrobacterium rhizogenes] dbj BAB16212.1 (AP002086) similar to jhp0928 gene in Helicobacter pylori [Agrobacterium rhizogenes] Length = 1693
SeqID 1314	SA-2300.3	Contig101 (14589-17354 p)	9	Identities = 58/176 (32%), Positives = 83/176 (46%), Gaps = 19/176 (10%) gb AAD33086.1 AF071083_1 (AF071083) fibronectin-binding protein I [Streptococcus pyogenes] Length = 1161
SeqID 1315	SA-2302.2	Contig79 (4055-5029 p)	67	Identities = 138/315 (43%), Positives = 222/315 (69%), Gaps = 6/315 (1%) pir C69763 ferrichrome ABC transporter (permease) homolog yclO - Bacillus subtilis dbj BAA09013.1 (D50453) homologue of ferric anguibactin transport system permease protein FatC of V. anguillarum [Bacillus subtilis] emb CAB12189.1 (Z99106) similar to ferrichrome ABC transporter (permease) [Bacillus subtilis] Length = 315

SeqID 1316	SA-2303.1	Contig79 (3096-4058 p)	67	<p>Identities = 149/304 (49%), Positives = 234/304 (76%) pir B69763 ferrichrome ABC transporter (permease) homolog yciN - Bacillus subtilis dbj BAA09012.1 (D50453) homologue of ferric anguibactin transport system permease protein FatD of V. anguillarum [Bacillus subtilis] emb CAB12188.1 (Z99106) similar to ferrichrome ABC transporter (permease) [Bacillus subtilis] Length = 316</p>
SeqID 1317	SA-2304.1	Contig79 (2309-2857 p)	63	<p>Identities = 93/182 (51%), Positives = 125/182 (68%), Gaps = 2/182 (1%) dbj BAB06720.1 (AP001517) maltose transacetylase (maltose O-acetyltransferase) [Bacillus halodurans] Length = 186</p>
SeqID 1318	SA-2305.1	Contig79 (1524-2288 p)	64	<p>Identities = 128/249 (51%), Positives = 168/249 (67%) sp O31744 RNH2_BACSU RIBONUCLEASE HII (RNASE HII) pir C69693 ribonuclease H rnh - Bacillus subtilis emb CAB13479.1 (Z99112) ribonuclease H [Bacillus subtilis] Length = 255</p>
SeqID 1319	SA-2306.1	Contig79 (689-1540 p)	66	<p>Identities = 141/281 (50%), Positives = 196/281 (69%), Gaps = 5/281 (1%) dbj BAA75361.1 (AB013365) YlqF [Bacillus halodurans] dbj BAB06195.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 284</p>
SeqID 1320	SA-2307.1	Contig79 (72-413 p)	62	<p>Identities = 61/135 (45%), Positives = 86/135 (63%), Gaps = 4/135 (2%) pir A69760 conserved hypothetical protein yciB - Bacillus subtilis dbj BAA08969.1 (D50453) yciB [Bacillus subtilis] emb CAB12129.1 (Z99105) similar to hypothetical proteins [Bacillus subtilis] Length = 194</p>
SeqID 1321	SA-2308.1	Contig74 (5434-6534 m)	40	<p>Identities = 85/336 (25%), Positives = 158/336 (46%), Gaps = 28/336 (8%) pir E71665 bicyclomycin resistance protein (bcr1) RP603 - Rickettsia prowazekii emb CAA15047.1 (AJ235272) BICYCLOMYCIN RESISTANCE PROTEIN (bcr1) [Rickettsia prowazekii] Length = 407</p>

SeqID 1322	SA-2309.1	Contig74 (4435-5382 m)	53	Identities = 127/269 (47%), Positives = 177/269 (65%), Gaps = 8/269 (2%) gb AAC23746.1 (AF052209) competence protein [Streptococcus pneumoniae] Length = 266
SeqID 1323	SA-231.1	Contig136 (35124-35423 m)	No Hits found	
SeqID 1324	SA-2311.1	Contig74 (2614-4419 m)	98	Identities = 593/601 (98%), Positives = 597/601 (98%) sp Q53778 PEPB_STRAG_GROUP_B_OLIGOPEPTIDASE_PEPB_pir T51748 thimet oligopeptidase (EC 3.4.24.15) PepB [validated] - Streptococcus agalactiae gb AAC44215.1 (U49821) group B oligopeptidase PepB [Streptococcus agalactiae] Length = 601
SeqID 1325	SA-2312.1	Contig74 (1793-2419 m)	28	Identities = 39/117 (33%), Positives = 67/117 (56%), Gaps = 9/117 (7%) pir A57362 gyrb protein - Streptococcus pneumoniae (fragment) emb CAA58770.1 (X83917) orfgyrb [Streptococcus pneumoniae] emb CAA91552.1 (Z67740) unidentified [Streptococcus pneumoniae] Length = 144
SeqID 1326	SA-2313.1	Contig74 (1012-1719 m)	69	Identities = 131/227 (57%), Positives = 169/227 (73%) emb CAA68045.1 (X99710) methyltransferase [Lactococcus lactis] Length = 227
SeqID 1327	SA-2314.1	Contig74 (22-951 m)	48	Identities = 101/307 (32%), Positives = 151/307 (48%), Gaps = 17/307 (5%) sp P15294 PRTM_LACLA_PROTEASE_MATURATION_PROTEIN_PRECURSOR_pir S08083 probable protein export protein prtM precursor - Lactococcus lactis subsp. cremoris (strain NCDO 763) plasmid pLP763 emb CAA32349.1 (X14130) ORF (AA 1 to 299) [Lactococcus lactis subsp. cremoris] Length = 299
SeqID 1328	SA-2315.2	Contig123 (8457-10655 m)	82	Identities = 539/725 (74%), Positives = 616/725 (84%), Gaps = 7/725 (0%) gb AAD00215.1 (U73336) anaerobic ribonucleotide reductase [Lactococcus lactis subsp. cremoris] Length = 747
SeqID 1329	SA-2316.1	Contig123 (8239-8382 m)	No Hits found	

SeqID 1330	SA-2317.1	Contig123 (7294-8226 m)	46	Identities = 91/299 (30%), Positives = 147/299 (48%), Gaps = 7/299 (2%) emb CAB95794.1 _[(AL359949) putative oxidoreductase [Streptomyces coelicolor A3(2)] Length = 301
SeqID 1331	SA-2318.1	Contig123 (6794-7285 m)	40	Identities = 52/129 (40%), Positives = 70/129 (53%), Gaps = 5/129 (3%) dbj BAB04222.1 (AP001508) unknown conserved protein in others [Bacillus halodurans] Length = 174
SeqID 1332	SA-2319.2	Contig123 (6104-6721 m)	79	Identities = 152/198 (76%), Positives = 176/198 (88%) gb AAD00216.1 (U73336) anaerobic ribonucleotide reductase activator protein [Lactococcus lactis subsp. cremoris] Length = 199
SeqID 1333	SA-232.1	Contig136 (34814-35113 m)	No Hits found	
SeqID 1334	SA-2322.2	Contig118 (25862-27211 p)	76	Identities = 275/450 (61%), Positives = 347/450 (77%), Gaps = 1/450 (0%) dbj BAA76640.1 (AB019579) glutathione reductase (GR) [Streptococcus mutans] Length = 450
SeqID 1335	SA-2324.1	Contig118 (27254-27706 m)	50	Identities = 45/156 (28%), Positives = 79/156 (49%), Gaps = 3/156 (1%) gb AAF87093.1 AF167576_1 (AF167576) secreted antigen SagBb [Enterococcus hirae] Length = 576
SeqID 1336	SA-2326.1	Contig118 (27953-29098 p)	60	Identities = 175/353 (49%), Positives = 234/353 (65%), Gaps = 1/353 (0%) sp P31672 NIFS_LACDE NIFS PROTEIN HOMOLOG pir S16047 nitrogenase cofactor synthesis protein nifs - Lactobacillus delbrueckii emb CAA43493.1 (X61190) nifs-like gene [Lactobacillus delbrueckii] Length = 355
SeqID 1337	SA-2327.2	Contig118 (29100-30314 p)	74	Identities = 264/385 (68%), Positives = 312/385 (80%) gb AAK04477.1 AE006275_1 (AE006275) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 406
SeqID 1338	SA-2329.3	Contig118 (30416-31594 p)	28	Identities = 68/224 (30%), Positives = 117/224 (51%), Gaps = 10/224 (4%) ref NP_053211.1 pXO2-56 [Bacillus anthracis] gb AAF13661.1 AF188935_59 (AF188935) pXO2-56 [Bacillus anthracis] Length = 411

SeqID 1339	SA-233.1	Contig136 (34011-34712 m)	No Hits found	
SeqID 1340	SA-2330.1	Contig73 (1490-2371 m)	93	Identities = 260/293 (88%), Positives = 276/293 (93%) dbj BAB16889.1 (AB050113) class-II aldolase [Streptococcus bovis] Length = 293
SeqID 1341	SA-2331.1	Contig73 (488-1405 p)	60	Identities = 175/306 (57%), Positives = 220/306 (71%), Gaps = 3/306 (0%) sp P14295 DHL2_LACCO L-2-HYDROXYISOCAPROATE DEHYDROGENASE (L-HICDH) pir JJQ0114 L-2-hydroxyisocaproate dehydrogenase (EC 1.1.1.-) - Lactobacillus confusus gb AAA88213.1 (M31425) L-2-hydroxyisocaproate dehydrogenase [Weissella confusa] Length = 310
SeqID 1342	SA-2332.1	Contig73 (61-249 m)	72	Identities = 45/62 (72%), Positives = 53/62 (84%) sp P37807 RL28_BACSU 50S RIBOSOMAL PROTEIN L28 pir S39982 ribosomal protein L28 (rpmB) - Bacillus subtilis gb AAC36810.1 (L12244) ribosomal protein L28 [Bacillus subtilis] emb CAA74255.1 (Y13937) putative RpmB protein [Bacillus subtilis] emb CAB13455.1 (Z99112) ribosomal protein L28 [Bacillus subtilis] Length = 62
SeqID 1343	SA-2334.1	Contig101 (17522-19501 p)	13	Identities = 65/236 (27%), Positives = 92/236 (38%), Gaps = 47/236 (19%) pir S52348 hypothetical protein 2 - Lactobacillus leichmannii emb CAA57459.1 (X81869) orf2 [Lactobacillus leichmannii] Length = 507
SeqID 1344	SA-2335.2	Contig101 (19712-20491 p)	50	Identities = 91/298 (30%), Positives = 155/298 (51%), Gaps = 13/298 (4%) gb AAK04857.1 AE006309_6 (AE006309) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 432
SeqID 1345	SA-2336.1	Contig101 (20491-21375 p)	46	Identities = 95/271 (35%), Positives = 139/271 (51%), Gaps = 16/271 (5%) gb AAC13546.1 (AF019629) putative fimbria-associated protein [Actinomyces naeslundii] Length = 365
SeqID 1346	SA-2337.1	Contig101 (21423-22298 p)	30	Identities = 62/245 (25%), Positives = 96/245 (38%), Gaps = 45/245 (18%) dbj BAB04080.1 (AP001508) unknown [Bacillus halodurans] Length = 1661

SeqID 1347	SA-2339.2	Contig135 (2828-3526 p)	95	Identities = 228/232 (98%), Positives = 229/232 (98%), Gaps = 3/232 (1%) dbj BAA82278.1 (AB028896) CpslA [Streptococcus agalactiae] Length = 229
SeqID 1348	SA-234.1	Contig136 (32885-33970 m)	38	Identities = 103/342 (30%), Positives = 155/342 (45%), Gaps = 50/342 (14%) ref NP_053232.1 pXO2-78 [Bacillus anthracis] gb AAAF13682.1 AF188935_80 (AF188935) pXO2-78 [Bacillus anthracis] Length = 344
SeqID 1349	SA-2340.2	Contig135 (3539-4927 p)	95	Identities = 448/449 (99%), Positives = 448/449 (99%) pir T44643 galactosyl transferase cpsD [imported] - Streptococcus agalactiae gb AAD53066.1 AF163833_6 (AF163833) CpsE [Streptococcus agalactiae] Length = 449
SeqID 1350	SA-2341.1	Contig135 (4951-5400 p)	93	Identities = 149/149 (100%), Positives = 149/149 (100%) pir T44644 glycosyl transferase activity enhancer cpsG [imported] - Streptococcus agalactiae gb AAD53067.1 AF163833_7 (AF163833) CpsF [Streptococcus agalactiae] Length = 149
SeqID 1351	SA-2342.1	Contig135 (5400-5873 p)	94	Identities = 155/157 (98%), Positives = 155/157 (98%) pir T44645 glycosyl transferase cpsH [imported] - Streptococcus agalactiae gb AAD53068.1 AF163833_8 (AF163833) CpsG [Streptococcus agalactiae] Length = 157
SeqID 1352	SA-2343.1	Contig135 (5870-7015 p)	99	Identities = 380/381 (99%), Positives = 380/381 (99%) pir T44646 capsular polysaccharide repeating unit polymerase cpsI [imported] - Streptococcus agalactiae gb AAD53069.1 AF163833_9 (AF163833) CpsH [Streptococcus agalactiae] Length = 381
SeqID 1353	SA-2344.1	Contig135 (7012-7980 p)	98	Identities = 318/322 (98%), Positives = 320/322 (98%) pir T44647 glycosyl transferase cpsJ [imported] - Streptococcus agalactiae gb AAD53070.1 AF163833_10 (AF163833) CpsI [Streptococcus agalactiae] Length = 322

SeqID 1354	SA-2345.1	Contig135 (8014-8961 p)	96	Identities = 314/315 (99%), Positives = 315/315 (99%) dbj BAA33750.1 (AB017355) galactosyltransferase [Streptococcus agalactiae] dbj BAA82284.1 (AB028896) CpslaJ [Streptococcus agalactiae] Length = 315
SeqID 1355	SA-2346.2	Contig135 (9045-10001 p)	98	Identities = 318/318 (100%), Positives = 318/318 (100%) dbj BAA33751.1 (AB017355) cpsJ [Streptococcus agalactiae] dbj BAA82285.1 (AB028896) CpslaK [Streptococcus agalactiae] Length = 318
SeqID 1356	SA-2347.1	Contig75 (158-718 p)	51	Identities = 64/234 (27%), Positives = 122/234 (51%), Gaps = 3/234 (1%) sp O32095 YUEF_BACSU HYPOTHETICAL 40.9 KDA PROTEIN IN DEQ-ALD INTERGENIC REGION pir G70007 conserved hypothetical protein yueF - Bacillus subtilis emb CAB15168.1 (Z99120) similar to hypothetical proteins [Bacillus subtilis] Length = 369
SeqID 1357	SA-2348.1	Contig75 (814-1494 p)	57	Identities = 90/210 (42%), Positives = 136/210 (63%) sp Q02170 RADC_BACSU DNA REPAIR PROTEIN RADC HOMOLOG (ORFB) pir B45239 DNA repair protein homolog ysxA - Bacillus subtilis gb AAA22396.1 (M96343) homologous to E. coli radC gene product and to unidentified protein from Staphylococcus aureus [Bacillus subtilis] gb AAA22583.1 (L08793) putative [Bacillus subtilis] emb CAB14764.1 (Z99118) similar to DNA repair protein [Bacillus subtilis] Length = 231
SeqID 1358	SA-2349.1	Contig75 (1507-2145 m)	64	Identities = 96/202 (47%), Positives = 140/202 (68%), Gaps = 4/202 (1%) pir A69787 hypothetical protein ydiH - Bacillus subtilis dbj BAA19721.1 (D88802) ydiH [Bacillus subtilis] emb CAB12416.1 (Z99107) ydiH [Bacillus subtilis] Length = 215
SeqID 1359	SA-235.1	Contig136 (32600-32830 m)	No Hits found	
SeqID 1360	SA-2350.1	Contig75 (2300-2647 m)	No Hits found	

SeqID 1361	SA-2351.1	Contig75 (2649-3767 m)	63	Identities = 173/373 (46%), Positives = 236/373 (62%), Gaps = 6/373 (1%) dbj BAB04979.1 (AP001511) Fe-S cluster formation protein [Bacillus halodurans] Length = 386
SeqID 1362	SA-2352.1	Contig75 (3768-4043 m)	68	Identities = 42/89 (47%), Positives = 63/89 (70%), Gaps = 2/89 (2%) gb AAF15359.1 AF201954_1 (AF201954) phosphoribosylpyrophosphate synthetase [Plasmodium falciparum] Length = 323
SeqID 1363	SA-2353.1	Contig72 (5600-5809 p)	82	Identities = 48/70 (68%), Positives = 58/70 (82%) pir T00087 rhamnosyltransferase - Streptococcus mutans dbj BAA32090.1 (AB010970) rhamnosyltransferase [Streptococcus mutans] Length = 311
SeqID 1364	SA-2354.1	Contig72 (4453-5610 p)	66	Identities = 234/362 (64%), Positives = 284/362 (77%) pir T00086 rgpAc protein - Streptococcus mutans dbj BAA32089.1 (AB010970) rgpAc [Streptococcus mutans] Length = 362
SeqID 1365	SA-2355.1	Contig72 (3485-4339 p)	92	Identities = 257/283 (90%), Positives = 273/283 (95%) gb AAC38675.1 (AF030359) dTDP-L-rhamnose synthase [Streptococcus pneumoniae] gb AAC38685.1 (AF030361) dTDP-L-rhamnose synthase [Streptococcus pneumoniae] gb AAC38701.1 (AF030364) dTDP-L-rhamnose synthase [Streptococcus pneumoniae] gb AAD10184.1 (AF026471) Cps2O [Streptococcus pneumoniae] Length = 283
SeqID 1366	SA-2356.1	Contig72 (3054-3395 p)	82	Identities = 92/108 (85%), Positives = 100/108 (92%) dbj BAA21508.1 (AB000631) unnamed protein product [Streptococcus mutans] Length = 111
SeqID 1367	SA-2357.1	Contig72 (1803-2945 p)	93	Identities = 345/367 (94%), Positives = 358/367 (97%) dbj BAA21507.1 (AB000631) sigma 42 protein [Streptococcus mutans] Length = 371
SeqID 1368	SA-2358.1	Contig72 (20-1828 p)	64	Identities = 271/637 (42%), Positives = 389/637 (60%), Gaps = 56/637 (8%) sp Q04505 PRIM_LACLA DNA PRIMASE pir JC2485 DNA primase (EC 2.7.7.-) dnaG - Lactococcus lactis prf 2106154A DNA primase [Lactococcus lactis] Length = 642

SeqID 1369	SA-2359.2	Contig71 (2968-3543 m)	63	Identities = 96/195 (49%), Positives = 138/195 (70%), Gaps = 15/195 (7%) gb AAK04732.1 AE006296_6 (AE006296) DNA-directed RNA polymerase delta chain (EC 2.7.7.6) [Lactococcus lactis subsp. lactis] Length = 187
SeqID 1370	SA-236.1	Contig136 (32214-32603 m)	No Hits found	
SeqID 1371	SA-2360.1	Contig71 (1091-2695 m)	89	Identities = 421/535 (78%), Positives = 481/535 (89%) emb CAA09021.2 (AJ010153) CTP synthetase [Lactococcus lactis subsp. cremoris] Length = 535
SeqID 1372	SA-2361.2	Contig71 (56-982 m)	67	Identities = 157/312 (50%), Positives = 212/312 (67%), Gaps = 9/312 (2%) gb AAK04219.1 AE006250_6 (AE006250) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 311
SeqID 1373	SA-2363.1	Contig104 (71-529 m)	No Hits found	
SeqID 1374	SA-2364.1	Contig104 (3-842 p)	38	Identities = 62/235 (26%), Positives = 108/235 (45%), Gaps = 12/235 (5%) pir T34651 probable transmembrane protein - Streptomyces coelicolor emb CAA22372.1 (AL034446) putative transmembrane protein [Streptomyces coelicolor A3(2)] Length = 307
SeqID 1375	SA-2365.1	Contig104 (949-1539 p)	No Hits found	
SeqID 1376	SA-2366.1	Contig104 (1573-2844 p)	47	Identities = 106/440 (24%), Positives = 206/440 (46%), Gaps = 65/440 (14%) pir C82901 conserved hypothetical UU367 [imported] - Ureaplasma urealyticum gb AAF30776.1 AE002133_9 (AE002133) conserved hypothetical [Ureaplasma urealyticum] Length = 507
SeqID 1377	SA-2367.1	Contig104 (2857-3288 p)	45	Identities = 39/145 (26%), Positives = 68/145 (46%), Gaps = 7/145 (4%) pir D69831 conserved hypothetical protein yhfO - Bacillus subtilis emb CAA74538.1 (Y14084) hypothetical protein [Bacillus subtilis] emb CAB12871.1 (Z99109) similar to hypothetical proteins [Bacillus subtilis] Length = 149

SeqID 1378	SA-2368.1	Contig104 (3380-4264 p)	64	Identities = 145/283 (51%), Positives = 191/283 (67%), Gaps = 12/283 (4%) dbj BAB06129.1 (AP001515) tRNA pseudouridine 5S synthase [Bacillus halodurans] Length = 304
SeqID 1379	SA-2369.2	Contig104 (4277-5209 p)	98	Identities = 310/311 (99%), Positives = 311/311 (99%) gb AAB64408.1 (U92073) macrolide-efflux protein [Streptococcus agalactiae] Length = 311
SeqID 1380	SA-237.1	Contig136 (31877-32062 m)	No Hits found	
SeqID 1381	SA-2370.3	Contig124 (2287-4089 m)	59	Identities = 252/598 (42%), Positives = 358/598 (59%), Gaps = 54/598 (9%) gb AAK04733.1 AE006296_7 (AE006296) conserved hypothetical protein [Lactococcus lactis subsp. lactis] Length = 550
SeqID 1382	SA-2371.1	Contig124 (1732-2214 m)	68	Identities = 86/154 (55%), Positives = 114/154 (73%), Gaps = 1/154 (0%) sp P80240 GRE_A_BACSU TRANSCRIPTION ELONGATION FACTOR GRE_A (TRANSCRIPT CLEAVAGE FACTOR GRE_A) (GENERAL STRESS PROTEIN 20M) (GSP20M) pir A69637 transcription elongation factor greA - Bacillus subtilis emb CAB14674.1 (Z99117) transcription elongation factor [Bacillus subtilis] Length = 157
SeqID 1383	SA-2372.1	Contig124 (168-1628 m)	54	Identities = 182/488 (37%), Positives = 267/488 (54%), Gaps = 17/488 (3%) pir C75543 6-aminohexanoate-cyclic-dimer hydrolase - Deinococcus radiodurans (strain R1) gb AAF09821.1 AE001885_5 (AE001885) 6-aminohexanoate-cyclic-dimer hydrolase [Deinococcus radiodurans] Length = 561
SeqID 1384	SA-2373.1	Contig93 (16319-16855 p)	54	Identities = 65/206 (31%), Positives = 110/206 (52%), Gaps = 16/206 (7%) gb AAA27630.1 (M98350) hydrophobic protein [unidentified bacterium] Length = 251

SeqID 1385	SA-2374.1	Contig93 (15061-16188 m)	22	<p>Identities = 54/192 (28%), Positives = 89/192 (46%), Gaps = 14/192 (7%) sp P50736 YPDA_BACSU HYPOTHETICAL 36.3 KD PROTEIN IN RECQ-CMK INTERGENIC REGION</p> <p>pir A69934 thioredoxin reductase homolog ypdA - Bacillus subtilis gb AAC83954.1 (L47648) putative [Bacillus subtilis] emb CAB14211.1 (Z99115) similar to thioredoxin reductase [Bacillus subtilis] emb CAB14227.1 (Z99116) similar to thioredoxin reductase [Bacillus subtilis] Length = 324</p>
SeqID 1386	SA-2376.1	Contig93 (14038-15024 p)	78	<p>Identities = 213/322 (66%), Positives = 260/322 (80%), Gaps = 5/322 (1%) gb AAB81912.1 (U92974) unknown [Lactococcus lactis] Length = 319</p>
SeqID 1387	SA-2377.1	Contig93 (13570-13875 m)	No Hits found	
SeqID 1388	SA-2378.2	Contig93 (12039-13412 p)	52	<p>Identities = 174/321 (54%), Positives = 241/321 (74%), Gaps = 4/321 (1%) pir JC5050 sugar phosphate transport protein - Shigella flexneri gb AAC44575.1 (U28354) IS629 ORFB fused with sequences similar to E. coli GlpT and UhpT proteins, Swiss-Prot Accession Number P08194 and P09836; Method: conceptual translation supplied by author [Shigella flexneri] Length = 333</p>
SeqID 1389	SA-238.1	Contig136 (31331-31807 m)	19	<p>Identities = 23/57 (40%), Positives = 36/57 (62%) ref NP_049417.1 putative cro-like regulatory protein [Streptococcus thermophilus bacteriophage DT1] gb AAD21905.1 (AF085222) putative cro-like regulatory protein [Streptococcus thermophilus bacteriophage DT1] Length = 67</p>
SeqID 1390	SA-2380.2	Contig130 (2285-3055 p)	55	<p>Identities = 112/253 (44%), Positives = 161/253 (63%), Gaps = 1/253 (0%) sp Q9RGS6 THIM_STACA HYDROXYETHYLTHIAZOLE KINASE (4-METHYL-5-BETA-HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE) gb AAF25543.1 (AF109218_3 (AF109218) ThiM [Staphylococcus carnosus] Length = 264</p>

SeqID 1391	SA-2381.1	Contig130 (1486-2283 p)	- - - 67	Identities = 139/258 (53%), Positives = 186/258 (71%), Gaps = 4/258 (1%) gb AAAF25542.1 AF109218_2 (AF109218) ThID [Staphylococcus carnosus] Length = 273
SeqID 1392	SA-2382.1	Contig130 (805-1461 p)	52	Identities = 74/213 (34%), Positives = 122/213 (56%), Gaps = 13/213 (6%) gb AAAF25541.1 AF109218_1 (AF109218) TenA [Staphylococcus carnosus] Length = 228
SeqID 1393	SA-2383.1	Contig130 (181-306 p)	No Hits found	
SeqID 1394	SA-2384.1	Contig130 (20-721 p)	41	Identities = 48/216 (22%), Positives = 98/216 (45%), Gaps = 3/216 (1%) emb CAA91230.1 (Z56283) orf2 [Lactobacillus helveticus] Length = 217
SeqID 1395	SA-2386.1	Contig101 (9386-10810 p)	64	Identities = 189/462 (40%), Positives = 313/462 (66%) emb CAB52225.1 (Z98171) EpsU protein [Streptococcus thermophilus] Length = 471
SeqID 1396	SA-2387.1	Contig101 (8022-9386 p)	No Hits found	
SeqID 1397	SA-2388.2	Contig101 (7073-8020 p)	23	Identities = 53/116 (45%), Positives = 75/116 (63%), Gaps = 4/116 (3%) gb AAF18951.1 AF155805_5 (AF155805) Cps9H [Streptococcus suis] Length = 143
SeqID 1398	SA-239.1	Contig136 (30561-31331 m)	58	Identities = 99/260 (38%), Positives = 157/260 (60%), Gaps = 3/260 (1%) pir S45085 hypothetical protein zeta - Streptococcus pyogenes plasmid pDB101 and pBT233 pir S68606 hypothetical protein zeta - Streptococcus pyogenes plasmid pBT233 emb CAA45934.1 (X64695) ORF zeta [Streptococcus pyogenes] emb CAA47091.1 (X66468) orf zeta [Streptococcus pyogenes] emb CAA47092.1 (X66468) orf zeta [Streptococcus pyogenes] Length = 287

SeqID 1399	SA-2390.2	Contig136 (91497-93569 p)	38	<p>Identities = 178/535 (33%), Positives = 269/535 (50%), Gaps = 55/535 (10%) sp P54602 YHCR_BACSU HYPOTHETICAL 132.7 KDA PROTEIN IN CSPB-GLPP INTERGENIC REGION pir F69823 probable phosphoesterase (EC 3.1.-.-) yhcR - Bacillus subtilis emb CAA65702.1 (X96983) hypothetical protein [Bacillus subtilis] emb CAB12747.1 (Z99108) similar to 5 - nucleotidase [Bacillus subtilis] Length = 1217</p>
SeqID 1400	SA-2391.1	Contig136 (93606-94016 m)	70	<p>Identities = 72/136 (52%), Positives = 96/136 (69%) sp O08450 DEF_CLOBE POLYPEPTIDE DEFORMYLASE (PDF) (FORMYLMETHIONINE DEFORMYLASE) emb CAB09662.1 (Z96934) peptide deformylase [Clostridium beijerinckii] Length = 136</p>
SeqID 1401	SA-2392.1	Contig136 (94086-95393 m)	81	<p>Identities = 292/436 (66%), Positives = 356/436 (80%), Gaps = 2/436 (0%) dbj BAB05820.1 (AP001514) NADP-specific glutamate dehydrogenase [Bacillus halodurans] Length = 458</p>
SeqID 1402	SA-2394.2	Contig89 (9769-11001 m)	29	<p>Identities = 63/243 (25%), Positives = 120/243 (48%) dbj BAB03800.1 (AP001507) BH0081~unknown conserved protein in others [Bacillus halodurans] Length = 251</p>
SeqID 1403	SA-2395.1	Contig89 (11041-12582 m)	81	<p>Identities = 352/509 (69%), Positives = 421/509 (82%), Gaps = 1/509 (0%) pir D69813 ABC transporter (ATP-binding protein) homolog yfmM - Bacillus subtilis dbj BAA22327.1 (D86417) YfmM [Bacillus subtilis] emb CAB12571.1 (Z99108) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 518</p>
SeqID 1404	SA-2396.2	Contig88 (2596-3666 p)	60	<p>Identities = 138/358 (38%), Positives = 220/358 (60%), Gaps = 18/358 (5%) pir E71373 probable regulatory protein (pfoS/R) - syphilis spirochete gb AAC65034.1 (AE001189) regulatory protein (pfoS/R) [Treponema pallidum] Length = 350</p>

SeqID 1405	SA-2397.1	Contig88 (1467-2459 p)	67	Identities = 163/325 (50%), Positives = 222/325 (68%), Gaps = 3/325 (0%) dbj BAB07127.1 (AP001518) thioredoxin reductase [Bacillus halodurans] Length = 330
SeqID 1406	SA-2398.1	Contig88 (731-1486 p)	73	Identities = 144/246 (58%), Positives = 186/246 (75%), Gaps = 6/246 (2%) dbj BAB06198.1 (AP001515) tRNA methyltransferase [Bacillus halodurans] Length = 246
SeqID 1407	SA-2399.1	Contig88 (1-744 p)	51	Identities = 88/174 (50%), Positives = 128/174 (72%), Gaps = 1/174 (0%) sp O31740 RIMM_BACSU PROBABLE 16S RRNA PROCESSING PROTEIN RIMM pir E69880 conserved hypothetical protein yiqE - Bacillus subtilis emb CAB13475.1 (Z99112) similar to hypothetical proteins [Bacillus subtilis] Length = 174
SeqID 1408	SA-240.2	Contig136 (28308-30458 m)	15	Identities = 65/236 (27%), Positives = 117/236 (49%), Gaps = 27/236 (11%) pir H83403 hypothetical protein PA1939 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG05327.1 AE004620_6 (AE004620) hypothetical protein [Pseudomonas aeruginosa] Length = 665
SeqID 1409	SA-2404.1	Contig122 (33097-35049 m)	77	Identities = 412/632 (65%), Positives = 506/632 (79%), Gaps = 6/632 (0%) pir S68599 phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus sobrinus (strain 6715) Length = 632
SeqID 1410	SA-2405.1	Contig70 (5129-5893 m)	98	Identities = 252/254 (99%), Positives = 253/254 (99%) gb AAG09975.1 AF248038_4 (AF248038) methyltransferase [Streptococcus agalactiae] Length = 254
SeqID 1411	SA-2406.1	Contig70 (4770-5108 m)	95	Identities = 112/112 (100%), Positives = 112/112 (100%) gb AAG09974.1 AF248038_3 (AF248038) unknown [Streptococcus agalactiae] Length = 112
SeqID 1412	SA-2407.1	Contig70 (4463-4768 m)	94	Identities = 101/101 (100%), Positives = 101/101 (100%) gb AAG09973.1 AF248038_2 (AF248038) acetate kinase [Streptococcus agalactiae] Length = 101

SeqID 1413	SA-2409.2	Contig70 (3706-4245 p)	52	Identities = 105/107 (98%), Positives = 106/107 (98%) gb AAG09972.1 AF248038_1 (AF248038) integrase [Streptococcus agalactiae] Length = 108
SeqID 1414	SA-241.2	Contig138 (37833-38351 m)	55	Identities = 74/175 (42%), Positives = 98/175 (55%), Gaps = 6/175 (3%) sp P50838 YPSA_BACSU HYPOTHETICAL 21.1 KD PROTEIN IN COTD-KDUD INTERGENIC REGION pir D69941 conserved hypothetical protein ypsA - Bacillus subtilis gb AAB38471.1 (L47838) putative [Bacillus subtilis] emb CAB14136.1 (Z99115) similar to hypothetical proteins from B. subtilis [Bacillus subtilis] Length = 180
SeqID 1415	SA-2410.1	Contig70 (3110-3544 p)	42	Identities = 40/123 (32%), Positives = 63/123 (50%), Gaps = 9/123 (7%) gb AAD00268.1 (U77495) putative integrase [Leuconostoc oenos bacteriophage 10MC] Length = 348
SeqID 1416	SA-2411.1	Contig70 (2734-3060 p)	33	Identities = 24/72 (33%), Positives = 38/72 (52%), Gaps = 2/72 (2%) dbj BAB07266.1 (AP001519) unknown conserved protein in others [Bacillus halodurans] Length = 79
SeqID 1417	SA-2412.1	Contig70 (2007-2561 p)	No Hits found	
SeqID 1418	SA-2414.1	Contig70 (1022-1783 p)	38	Identities = 70/211 (33%), Positives = 101/211 (47%), Gaps = 5/211 (2%) pir JH0204 hypothetical 30.5K protein precursor - Enterococcus faecalis plasmid pAM-beta-1 gb AAC38600.1 (AF007787) orfC [Enterococcus faecalis] Length = 288
SeqID 1419	SA-2415.1	Contig70 (440-997 p)	No Hits found	
SeqID 1420	SA-2416.1	Contig70 (12-440 p)	No Hits found	
SeqID 1421	SA-2418.2	Contig104 (16088-16768 p)	88	Identities = 202/246 (82%), Positives = 223/246 (90%), Gaps = 7/246 (2%) emb CAB54564.1 (AJ005926) response regulator [Streptococcus pneumoniae] Length = 246
SeqID 1422	SA-2419.2	Contig104 (13377-15926 p)	76	Identities = 555/847 (65%), Positives = 673/847 (78%), Gaps = 4/847 (0%) emb CAB50785.1 (AJ007700) aminopeptidase N [Streptococcus thermophilus] Length = 847

SeqID 1423	SA-242.2	Contig138 (38426-39025 p)	78	Identities = 143/196 (72%), Positives = 165/196 (83%), Gaps = 1/196 (0%) sp Q00579 RECU_STOR RECOMBINATION PROTEIN U HOMOLOG gb AAA26957.1 (M90528) ORF [Streptococcus oralis] Length = 198
SeqID 1424	SA-2420.2	Contig118 (31784-32098 p)	70	Identities = 82/104 (78%), Positives = 92/104 (87%) gb AAK05177.1 AE006340_10 (AE006340) 50S ribosomal protein L21 [Lactococcus lactis subsp. lactis] Length = 104
SeqID 1425	SA-2421.2	Contig118 (31661-32134 m)	No Hits found	
SeqID 1426	SA-2422.2	Contig118 (32105-32443 p)	53	Identities = 38/107 (35%), Positives = 61/107 (56%), Gaps = 5/107 (4%) gb AAK05178.1 AE006340_11 (AE006340) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 111
SeqID 1427	SA-2423.1	Contig118 (32465-32758 p)	76	Identities = 70/90 (77%), Positives = 80/90 (88%) sp P05657 RL27_BACSU 50S RIBOSOMAL PROTEIN L27 (BL30) (BL24) pir C21895 ribosomal protein L27 - Bacillus subtilis emb CAA26492.1 (X02656) homologous to E.coli ribosomal protein L27 [Bacillus subtilis] emb CAB14754.1 (Z99118) ribosomal protein L27 (BL24) [Bacillus subtilis] Length = 94
SeqID 1428	SA-2424.1	Contig118 (32974-33879 p)	53	Identities = 105/297 (35%), Positives = 164/297 (54%), Gaps = 4/297 (1%) pir T44638 capsular polysaccharide biosynthesis protein cpsY [imported] - Streptococcus agalactiae emb CAB36980.1 (Y17218) CpsY protein [Streptococcus agalactiae] emb CAB36982.2 (Y17241) CpsY protein [Streptococcus agalactiae] gb AAD53064.1 AF163833_1 (AF163833) CpsY [Streptococcus agalactiae] Length = 307
SeqID 1429	SA-2425.1	Contig118 (33888-34352 p)	58	Identities = 61/144 (42%), Positives = 94/144 (64%), Gaps = 1/144 (0%) sp Q48729 LSPA_LACLC LIPOPROTEIN SIGNAL PEPTIDASE (PROLIPOPROTEIN SIGNAL PEPTIDASE) (SIGNAL PEPTIDASE II) (SPASE II) Length = 150

SeqID 1430	SA-2426.2	Contig118 (34336-34836 p)	80	Identities = 115/168 (68%), Positives = 140/168 (82%) gb AAK05096.1 AE006334_2 (AE006334) pseudouridine synthase [Lactococcus lactis subsp. lactis] Length = 301
SeqID 1431	SA-2428.1	Contig67 (20-2098 p)	87	Identities = 536/692 (77%), Positives = 613/692 (88%) emb CAC09927.1 (AJ249559) translation elongation factor G, EF- G [Bacillus stearothermophilus] Length = 692
SeqID 1432	SA-2429.2	Contig135 (26612-27559 m)	23	Identities = 42/173 (24%), Positives = 80/173 (45%), Gaps = 2/173 (1%) ref NP_052734.1 pXO1-38 [Bacillus anthracis] pir F59095 hypothetical protein pXO1-38 - Bacillus anthracis virulence plasmid pXO1 gb AAD32342.1 AAD32342 (AF065404) pXO1-38 [Bacillus anthracis] Length = 276
SeqID 1433	SA-2430.1	Contig135 (25417-26493 m)	64	Identities = 153/350 (43%), Positives = 234/350 (66%), Gaps = 4/350 (1%) ref NP_049990.1 orf359 gp [Streptococcus thermophilus bacteriophage Sfi21] emb CAA64931.1 (X95646) integrase [Streptococcus thermophilus bacteriophage Sfi21] gb AAC03454.1 (AF020798) integrase homolog [Streptococcus thermophilus bacteriophage TP-J34] gb AAD44095.1 AF115103_25 (AF115103) orf359 gp [Streptococcus thermophilus bacteriophage Sfi21] Length = 359
SeqID 1434	SA-2431.3	Contig135 (23774-25237 p)	63	Identities = 251/401 (62%), Positives = 312/401 (77%), Gaps = 4/401 (0%) gb AAK04936.1 AE006317_5 (AE006317) 30S ribosomal protein S1 [Lactococcus lactis subsp. lactis] Length = 408

SeqID 1435	SA-2432.3	Contig135 (24256-25128 m)	38	Identities = 73/265 (27%), Positives = 111/265 (41%), Gaps = 40/265 (15%) ref NP_014442.1 anchorage subunit of a- agglutinin; Aga1p [Saccharomyces cerevisiae] sp P32323 AGA1_YEAST A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR pir A41258 a-agglutinin core protein AGA1 - yeast (Saccharomyces cerevisiae) gb AAA34382.1 (M60590) a-agglutinin core subunit [Saccharomyces cerevisiae] emb CAA96325.1 (Z71659) ORF YNR044w [Saccharomyces cerevisiae] Length = 725
SeqID 1436	SA-2433.2	Contig96 (1706-2137 m)	42	Identities = 47/137 (34%), Positives = 71/137 (51%), Gaps = 5/137 (3%) dbj BAB04953.1 (AP001511) small multidrug export related protein [Bacillus halodurans] Length = 134
SeqID 1437	SA-2434.3	Contig96 (507-1547 p)	81	Identities = 242/338 (71%), Positives = 290/338 (85%) gb AAK06250.1 AE006444_1 (AE006444) elongation factor Ts [Lactococcus lactis subsp. lactis] Length = 342
SeqID 1438	SA-2435.3	Contig96 (3-413 p)	83	Identities = 102/131 (77%), Positives = 115/131 (86%), Gaps = 2/131 (1%) gb AAK06251.1 AE006444_2 (AE006444) 30S ribosomal protein S2 [Lactococcus lactis subsp. lactis] Length = 255
SeqID 1439	SA-2436.3	Contig96 (2-316 m)	30	Identities = 28/59 (47%), Positives = 32/59 (53%) emb CAB67155.1 (AJ271079) hypothetical protein [Oenothera elata subsp. hookeri] Length = 94
SeqID 1440	SA-2437.2	Contig135 (2125-2817 p)	93	Identities = 229/230 (99%), Positives = 229/230 (99%) sp Q04662 CPSB_STRAG CPSB PROTEIN pir S34975 polysaccharide chain length regulator cpsB [imported] - Streptococcus agalactiae gb AAB00362.1 (AF163833) CpsC [Streptococcus agalactiae] Length = 230
SeqID 1441	SA-2438.1	Contig135 (1385-2116 p)	96	Identities = 242/243 (99%), Positives = 243/243 (99%) dbj BAA82276.1 (AB028896) CpslaB [Streptococcus agalactiae] Length = 243
SeqID 1442	SA-2439.1	Contig135 (80-427 m)	No Hits found	

SeqID 1443	SA-244.2	Contig138 (39012-41258 p)	71	Identities = 412/725 (56%), Positives = 537/725 (73%), Gaps = 25/725 (3%) gb AAF17262.1 AF210752_1 (AF210752) penicillin-binding protein 1A [Streptococcus pneumoniae] Length = 719
SeqID 1444	SA-2440.1	Contig135 (12-1379 p)	99	Identities = 450/453 (99%), Positives = 452/453 (99%) emb CAB36981.1 (Y17218) CpsX protein [Streptococcus agalactiae] dbj BAA82275.1 (AB028896) CpsIA [Streptococcus agalactiae] Length = 485
SeqID 1445	SA-2441.1	Contig123 (43050-43538 m)	No Hits found	
SeqID 1446	SA-2442.1	Contig123 (42359-42814 p)	No Hits found	
SeqID 1447	SA-2443.1	Contig123 (41386-42402 p)	No Hits found	
SeqID 1448	SA-2445.2	Contig123 (39711-40976 p)	48	Identities = 138/398 (34%), Positives = 208/398 (51%), Gaps = 17/398 (4%) ref NP_076752.1 integrase [bacteriophage bIL310] gb AAK04145.1 AE006243_9 (AE006243) prophage ps1 protein 23, integrase [Lactococcus lactis subsp. lactis] gb AAK08405.1 AF323671_1 (AF323671) integrase [bacteriophage bIL310] Length = 394
SeqID 1449	SA-2446.2	Contig136 (2790-3302 m)	48	Identities = 78/191 (40%), Positives = 112/191 (57%), Gaps = 9/191 (4%) dbj BAA21095.1 (D88438) repB [Lactobacillus acidophilus] Length = 193
SeqID 1450	SA-2447.1	Contig136 (2553-2756 m)	No Hits found	
SeqID 1451	SA-2448.1	Contig136 (1354-2535 m)	41	Identities = 101/370 (27%), Positives = 167/370 (44%), Gaps = 39/370 (10%) pir T13289 probable integrase - Streptococcus phage phi-O1205 gb AAC79517.1 (U88974) ORF1 [Streptococcus thermophilus temperate bacteriophage O1205] Length = 359
SeqID 1452	SA-2450.2	Contig135 (78598-79290 m)	No Hits found	
SeqID 1453	SA-2452.1	Contig66 (2566-2703 m)	No Hits found	

SeqID 1454	SA-2453.1	Contig66 (1714-2415 m)	67	Identities = 112/244 (45%), Positives = 164/244 (66%), Gaps = 10/244 (4%) sp O31458 YBFT_BACSU HYPOTHETICAL 27.3 KD PROTEIN IN GLTP-CWLJ INTERGENIC REGION pir E69750 glucosamine-6-phosphate isomerase homolog ybft - Bacillus subtilis emb CAB12030.1 (Z99105) similar to glucosamine-6-phosphate isomerase [Bacillus subtilis] dbj BAA33133.1 (AB006424) ybft [Bacillus subtilis] Length = 249
SeqID 1455	SA-2454.1	Contig66 (686-1642 p)	70	Identities = 197/318 (61%), Positives = 243/318 (75%) pir S76960 hypothetical protein - Synechocystis sp. (strain PCC 6803) dbj BAA18872.1 (D90917) hypothetical protein [Synechocystis sp.] Length = 319
SeqID 1456	SA-2455.1	Contig66 (3-590 m)	60	Identities = 91/196 (46%), Positives = 120/196 (60%), Gaps = 1/196 (0%) dbj BAB06992.1 (AP001518) 16S pseudouridylylate synthase [Bacillus halodurans] Length = 238
SeqID 1457	SA-2457.3	Contig132 (33811-35607 p)	45	Identities = 151/620 (24%), Positives = 273/620 (43%), Gaps = 57/620 (9%) gb AAK06205.1 AE006439_2 (AE006439) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 605
SeqID 1458	SA-2458.3	Contig132 (33590-33832 p)	39	Identities = 33/83 (39%), Positives = 51/83 (60%) pir E83144 hypothetical protein PA4016 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG07403.1 AE004818_9 (AE004818) hypothetical protein [Pseudomonas aeruginosa] Length = 579
SeqID 1459	SA-2459.3	Contig91 (5194-5715 p)	61	Identities = 72/165 (43%), Positives = 112/165 (67%), Gaps = 2/165 (1%) sp P39667 YRXA_BACSU HYPOTHETICAL 19.7 KDA PROTEIN IN PHEA-NIFS INTERGENIC REGION (ORF1) pir A47071 hypothetical protein yrxA - Bacillus subtilis emb CAB14749.1 (Z99118) yrxA [Bacillus subtilis] Length = 178

SeqID 1460	SA-246.1	Contig138 (41304-42638 m)	90	Identities = 363/445 (81%), Positives = 408/445 (91%) sp Q56115 PEPC_STRTR AMINOPEPTIDASE C pir S48143 cysteine aminopeptidase C - Streptococcus thermophilus emb CAA82960.1 (Z30315) aminopeptidase C [Streptococcus thermophilus] Length = 445
SeqID 1461	SA-2460.1	Contig91 (4551-5126 p)	56	Identities = 97/188 (51%), Positives = 133/188 (70%) gb AAC18360.1 (AF064763) putative membrane spanning protein [Lactococcus lactis subsp. cremoris] Length = 196
SeqID 1462	SA-2461.1	Contig91 (3911-4414 p)	No Hits found	
SeqID 1463	SA-2462.1	Contig91 (3130-3873 p)	59	Identities = 106/246 (43%), Positives = 150/246 (60%), Gaps = 7/246 (2%) pir G69984 rRNA methylase homolog ysgA - Bacillus subtilis emb CAA99602.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14825.1 (Z99118) similar to rRNA methylase [Bacillus subtilis] Length = 248
SeqID 1464	SA-2463.1	Contig91 (2820-3377 m)	29	Identities = 35/91 (38%), Positives = 54/91 (58%), Gaps = 3/91 (3%) pir G72240 hypothetical protein TM1564 - Thermotoga maritima (strain MSB8) gb AAD36630.1 AE001801_17 (AE001801) acylphosphatase, putative [Thermotoga maritima] Length = 90
SeqID 1465	SA-2464.2	Contig91 (1803-2693 m)	63	Identities = 140/307 (45%), Positives = 201/307 (64%), Gaps = 19/307 (6%) gb AAK04667.1 AE006291_1 (AE006291) conserved hypothetical protein [Lactococcus lactis lactis] Length = 307 subsp.

SeqID 1466	SA-2465.1	Contig139 (189466-189870 m)	45	<p>Identities = 42/105 (40%), Positives = 62/105 (59%) sp P45678 PEB1_CAMJE MAJOR CELL-BINDING FACTOR PRECURSOR (CBF1) (PEB1) pir A48518 probable ABC-type amino-acid transporter periplasmic solute-binding protein Cj0921c precursor [imported] - Campylobacter jejuni (strain NCTC 11168) gb AAA02919.1 (L13662) major cell- binding factor [Campylobacter jejuni] emb CAB73178.1 (AL139076) probable ABC-type amino-acid transporter periplasmic solute-binding protein [Campylobacter jejuni] Length = 259</p>
SeqID 1467	SA-2466.1	Contig139 (188758-189453 m)	68	<p>Identities = 112/226 (49%), Positives = 161/226 (70%), Gaps = 3/226 (1%) pir G81365 probable ABC-type amino-acid transporter permease protein Cj0920c [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73177.1 (AL139076) putative ABC-type amino-acid transporter permease protein [Campylobacter jejuni] Length = 250</p>
SeqID 1468	SA-2467.1	Contig139 (188096-188746 m)	58	<p>Identities = 85/216 (39%), Positives = 132/216 (60%), Gaps = 6/216 (2%) pir F69633 glutamine ABC transporter (membrane protein) glhP - Bacillus subtilis emb CAB14687.1 (Z99117) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis] emb CAB14704.1 (Z99118) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis] Length = 218</p>
SeqID 1469	SA-2468.2	Contig139 (187481-188047 p)	42	<p>Identities = 43/157 (27%), Positives = 83/157 (52%), Gaps = 9/157 (5%) dbj BAB04094.1 (AP001508) BH0375~unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 194</p>
SeqID 1470	SA-2469.2	Contig139 (186295-187314 p)	No Hits found	

SeqID 1471	SA-247.1	Contig138 (42751-43572 m)	75	Identities = 173/275 (62%), Positives = 215/275 (77%), Gaps = 1/275 (0%) sp P18843 NADE_ECOLI_NH(3)-DEPENDENT NAD(+) SYNTHETASE (NITROGEN-REGULATORY PROTEIN) pir D64933 NAD+ synthase (EC 6.3.1.5) nadE [validated] - Escherichia coli (strain K-12) dbj BAA15529.1 (D90817) NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1) (Nitrogen-regulatory protein). [Escherichia coli] dbj BAA15535.1 (D90818) NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1) (Nitrogen-regulatory protein). [Escherichia coli] gb AAC74810.1 (AE000269) NAD synthetase, prefers NH3 over glutamine [Escherichia coli K12] Length = 275
SeqID 1472	SA-2470.1	Contig65 (2150-2422 m)	No Hits found	
SeqID 1473	SA-2471.1	Contig65 (1678-2109 p)	67	Identities = 69/152 (45%), Positives = 101/152 (66%), Gaps = 12/152 (7%) pir T11571 hypothetical protein 2 - Streptococcus mutans gb AAD15622.1 (U75480) unknown [Streptococcus mutans] Length = 151
SeqID 1474	SA-2472.1	Contig65 (1283-1681 p)	77	Identities = 88/129 (68%), Positives = 112/129 (86%) pir T11570 hypothetical protein 1 - Streptococcus mutans gb AAD15621.1 (U75480) unknown [Streptococcus mutans] Length = 131
SeqID 1475	SA-2473.1	Contig65 (495-1268 p)	85	Identities = 184/258 (71%), Positives = 227/258 (87%) sp P72482 LGT_STRMU_PROLIPOPROTEIN DIACYLGLYCERYL TRANSFERASE pir T11569 prolipoprotein diacylglycerol transferase (EC 2.4.99.-) - Streptococcus mutans gb AAC80171.3 (U75480) putative prolipoprotein diacylglycerol transferase [Streptococcus mutans] Length = 259
SeqID 1476	SA-2474.1	Contig65 (377-502 p)	89	Identities = 68/77 (88%), Positives = 75/77 (97%) sp Q9ZA98 HPRK_STRSL_HPR(SER) KINASE/PHOSPHATASE gb AAD12781.1 (AF069743) HPr(serine) kinase [Streptococcus salivarius] Length = 309

SeqID 1477	SA-2475.1	Contig65 (1-264 p)	95	Identities = 82/84 (97%), Positives = 84/84 (99%) dbj BAA7782.1 (AB027460) Hpr kinase [Streptococcus bovis] Length = 310
SeqID 1478	SA-2476.1	Contig64 (2858-3013 m)	75	Identities = 33/46 (71%), Positives = 39/46 (84%) pir H70958 probable uracil phosphoribosyltransferase (EC 2.4.2.9) - Mycobacterium tuberculosis (strain H37RV) emb CAB02640.1 (Z81011) pyrR [Mycobacterium tuberculosis] Length = 193
SeqID 1479	SA-2477.1	Contig64 (1739-2809 m)	72	Identities = 188/352 (53%), Positives = 265/352 (74%) emb CAB89872.1 (AJ132624) carbamoyl phosphate synthetase small subunit [Lactococcus lactis] Length = 357
SeqID 1480	SA-2479.1	Contig64 (85-1683 m)	37	Identities = 116/414 (28%), Positives = 204/414 (49%), Gaps = 31/414 (7%) sp P77886 CARB_LACPL CARBAMOYL- PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, LARGE CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE AMMONIA CHAIN) emb CAA91005.1 (Z54240) carbamoyl- phosphate synthase [Lactobacillus plantarum] Length = 1058
SeqID 1481	SA-248.1	Contig138 (43569-45029 m)	77	Identities = 309/476 (64%), Positives = 384/476 (79%), Gaps = 2/476 (0%) pir D70008 nicotinate phosphoribosyltransferase homolog yueK - Bacillus subtilis emb CAB15163.1 (Z99120) similar to nicotinate phosphoribosyltransferase [Bacillus subtilis] Length = 490
SeqID 1482	SA-2480.2	Contig79 (5026-5787 p)	60	Identities = 125/247 (50%), Positives = 187/247 (75%) pir D69763 ferrichrome ABC transporter (ATP-binding p) homolog yciP - Bacillus subtilis dbj BAA09014.1 (D50453) homologue of iron dicitrate transport ATP-binding protein FecE of E. coli [Bacillus subtilis] emb CAB12190.1 (Z99106) similar to ferrichrome ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 252

SeqID 1483	SA-2482.2	Contig79 (5849-6877 p)	55	Identities = 122/348 (35%), Positives = 201/348 (57%), Gaps = 16/348 (4%) emb CAA06500.1 (AJ005352) lipoprotein [Staphylococcus aureus] Length = 342
SeqID 1484	SA-2483.1	Contig79 (7015-7677 p)	44	Identities = 72/172 (41%), Positives = 108/172 (61%), Gaps = 2/172 (1%) pir C72399 DNA processing chain A - Thermotoga maritima (strain MSB8) gb AAD35341.1 AE001708_9 (AE001708) DNA processing chain A [Thermotoga maritima] Length = 337
SeqID 1485	SA-2486.2	Contig68 (423-1265 m)	64	Identities = 157/281 (55%), Positives = 196/281 (68%), Gaps = 6/281 (2%) gb AAK04366.1 AE006264_5 (AE006264) oxidoreductase [Lactococcus lactis subsp. lactis] Length = 281
SeqID 1486	SA-2488.2	Contig69 (4025-5050 m)	43	Identities = 96/352 (27%), Positives = 164/352 (46%), Gaps = 21/352 (5%) pir C69858 conserved hypothetical protein yknX - Bacillus subtilis emb CAB13308.1 (Z99111) similar to hypothetical proteins from B. subtilis [Bacillus subtilis] gb AAC24909.1 (AF012285) YknX [Bacillus subtilis] Length = 377
SeqID 1487	SA-249.1	Contig138 (45187-46101 m)	74	Identities = 173/302 (57%), Positives = 234/302 (77%) dbj BAB07290.1 (AP001519) thioredoxin reductase (NADPH) [Bacillus halodurans] Length = 315
SeqID 1488	SA-2490.2	Contig69 (3312-4022 m)	69	Identities = 131/218 (60%), Positives = 169/218 (77%) dbj BAB06841.1 (AP001517) ABC transporter (ATP-binding protein) [Bacillus halodurans] Length = 227
SeqID 1489	SA-2491.2	Contig69 (2050-3297 m)	57	Identities = 161/419 (38%), Positives = 243/419 (57%), Gaps = 25/419 (5%) sp O31712 YKNZ_BACSU HYPOTHETICAL 42.1 KDA PROTEIN IN MOAD-FRUR INTERGENIC REGION pir E69858 conserved hypothetical protein yknZ - Bacillus subtilis emb CAB13310.1 (Z99111) similar to hypothetical proteins [Bacillus subtilis] gb AAC24912.1 (AF012285) YknZ [Bacillus subtilis] Length = 397

SeqID 1490	SA-2492.2	Contig106 (19268-21082 p)	72	<p>Identities = 356/604 (58%), Positives = 445/604 (72%), Gaps = 4/604 (0%) sp P39754 GLMS_BACSU GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE [ISOMERIZING] (HEXOSEPHOSPHATE AMINOTRANSFERASE) (D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GFAT) (L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GLUCOSAMINE-6-PHOSPHATE SYNTHASE) pir B69633 glutamine--fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glmS - Bacillus subtilis gb AAA64224.1 (U21932) L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus subtilis] emb CAB11954.1 (Z99104) L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus subtilis] emb CAB11971.1 (Z99105) L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus subtilis] dbj BAA33071.1 (AB006424) L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDO TRANSFERASE [Bacillus subtilis] Length = 600</p>
SeqID 1491	SA-2493.1	Contig62 (1978-3021 p)	83	<p>Identities = 255/348 (73%), Positives = 298/348 (85%), Gaps = 1/348 (0%) sp P10539 DHAS_STRMU ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (ASA DEHYDROGENASE) (ASADH) pir A29137 aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) - Streptococcus mutans gb AAA26850.1 (J02667) aspartate beta-semialdehyde dehydrogenase (EC 1.2.1.11) [Streptococcus mutans] Length = 357</p>
SeqID 1492	SA-2494.1	Contig62 (640-1494 m)	34	<p>Identities = 71/271 (26%), Positives = 97/271 (35%), Gaps = 5/271 (1%) pir S54157 extensin-like protein - cowpea (fragment) Length = 279</p>
SeqID 1493	SA-2495.1	Contig62 (537-1769 p)	42	<p>Identities = 86/373 (23%), Positives = 176/373 (47%), Gaps = 26/373 (6%) gb AAF48863.1 (AE003509) CG15040 gene product [Drosophila melanogaster] Length = 1895</p>
SeqID 1494	SA-2497.2	Contig104 (5252-5665 p)	68	<p>Identities = 74/126 (58%), Positives = 101/126 (79%) gb AAF21893.1 AF103794_1 (AF103794) unknown [Listeria monocytogenes] Length = 131</p>

SeqID 1495	SA-2498.2	Contig104 (5658-5939 p)	47	Identities = 33/78 (42%), Positives = 50/78 (63%) pir C69864 hypothetical protein ykA - Bacillus subtilis emb CAB13337.1 (Z99111) ykA [Bacillus subtilis] gb AAC24938.1 (AF012285) unknown [Bacillus subtilis] Length = 88
SeqID 1496	SA-2499.3	Contig104 (5929-6693 p)	56	Identities = 121/252 (48%), Positives = 171/252 (67%), Gaps = 4/252 (1%) gb AAK04646.1 AE006288_9 (AE006288) conserved hypothetical protein [Lactococcus lactis subsp. lactis] Length = 255
SeqID 1497	SA-25.1	Contig137 (15915-20645 p)	11	Identities = 117/377 (31%), Positives = 182/377 (48%), Gaps = 73/377 (19%) ref NP_053235.1 pXO2-81 [Bacillus anthracis] gb AAF13685.1 AF188935_83 (AF188935) pXO2-81 [Bacillus anthracis] Length = 589
SeqID 1498	SA-250.1	Contig138 (46170-46418 m)	No Hits found	
SeqID 1499	SA-2500.3	Contig104 (6761-8071 p)	48	Identities = 139/466 (29%), Positives = 213/466 (44%), Gaps = 36/466 (7%) pir A82193 Sun/nucleolar protein family protein VC1502 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF94657.1 (AE004228) Sun/nucleolar protein family protein [Vibrio cholerae] Length = 503
SeqID 1500	SA-2501.2	Contig63 (2458-3795 m)	47	Identities = 111/428 (25%), Positives = 210/428 (48%), Gaps = 39/428 (9%) pir B69855 amino acid permease homolog ykba - Bacillus subtilis emb CAA05566.1 (AJ002571) Ykba [Bacillus subtilis] emb CAB13143.1 (Z99110) similar to amino acid permease [Bacillus subtilis] Length = 438
SeqID 1501	SA-2502.2	Contig63 (1737-2285 m)	71	Identities = 111/168 (66%), Positives = 136/168 (80%), Gaps = 3/168 (1%) gb AAK05247.1 AE006347_5 (AE006347) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 169

SeqID 1502	SA-2503.1	Contig63 (432-1601 m)	56	Identities = 130/381 (34%), Positives = 221/381 (57%), Gaps = 4/381 (1%) sp Q08432 PATB_BACSU PUTATIVE AMINOTRANSFERASE B pir S32934 aminotransferase patB - Bacillus subtilis emb CAB07910.1 (Z93933) aminotransferase [Bacillus subtilis] emb CAB07924.1 (Z93934) aminotransferase [Bacillus subtilis] gb AAB61979.1 (U63302) PatB [Bacillus subtilis] emb CAB15133.1 (Z99120) aminotransferase [Bacillus subtilis] Length = 387
SeqID 1503	SA-2504.1	Contig63 (179-322 m)	84	Identities = 44/47 (93%), Positives = 45/47 (95%) gb AAA71926.1 (L08445) uracil phosphoribosyltransferase [Streptococcus salivarius] Length = 128
SeqID 1504	SA-2505.1	Contig63 (3-182 m)	90	Identities = 49/60 (81%), Positives = 54/60 (89%) gb AAA71926.1 (L08445) uracil phosphoribosyltransferase [Streptococcus salivarius] Length = 128
SeqID 1505	SA-2507.2	Contig132 (1877-2116 p)	57	Identities = 27/75 (36%), Positives = 52/75 (69%) ref NP_050766.1 acyl carrier protein [Guillardia theta] sp P29189 ACP_GUITH ACYL CARRIER PROTEIN gb AAC35700.1 (AF041468) acyl carrier protein [Guillardia theta] Length = 81
SeqID 1506	SA-2508.1	Contig132 (874-1866 p)	70	Identities = 174/330 (52%), Positives = 239/330 (71%), Gaps = 2/330 (0%) sp P71018 PLSX_BACSU FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLSX pir H69679 involved in fatty acid/phospholipid synthesis plsX - Bacillus subtilis emb CAA74248.1 (Y13937) putative PLSX protein [Bacillus subtilis] emb CAB13462.1 (Z99112) alternate gene name: ylpD [Bacillus subtilis] Length = 333
SeqID 1507	SA-2509.1	Contig132 (2-796 p)	26	Identities = 47/154 (30%), Positives = 69/154 (44%), Gaps = 12/154 (7%) pir T34651 probable transmembrane protein - Streptomyces coelicolor emb CAA22372.1 (AL034446) putative transmembrane protein [Streptomyces coelicolor A3(2)] Length = 307

SeqID 1508	SA-251.1	Contig138 (46516-47259 m)	69	<p>Identities = 121/247 (48%), Positives = 176/247 (70%) sp O34900 YTMN_BACSU PROBABLE AMINO-ACID ABC TRANSPORTER ATP-BINDING PROTEIN YTMN pir F69641 histidine transport protein hisP - Bacillus subtilis gb AAC00329.1 (AF008220) putative amino acid transporter [Bacillus subtilis] emb CAB14894.1 (Z99118) histidine transport protein (ATP- binding protein) [Bacillus subtilis] emb CAB14912.1 (Z99119) histidine transport protein (ATP-binding protein) [Bacillus subtilis] Length = 259</p>
SeqID 1509	SA-2510.2	Contig135 (76527-78527 p)	79	<p>Identities = 430/659 (65%), Positives = 537/659 (81%), Gaps = 4/659 (0%) gb AAK04722.1 AE006295_3 (AE006295) potassium uptake protein [Lactococcus lactis subsp. lactis] Length = 671</p>
SeqID 1510	SA-2511.1	Contig135 (75631-76392 p)	67	<p>Identities = 142/248 (57%), Positives = 179/248 (71%), Gaps = 1/248 (0%) sp P45200 YDFG_HAEIN HYPOTHETICAL OXIDOREDUCTASE HI1430 pir H64122 ydfG protein - Haemophilus influenzae (strain Rd KW20) gb AAC23077.1 (U32822) short chain dehydrogenase/reductase [Haemophilus influenzae Rd] Length = 252</p>
SeqID 1511	SA-2512.2	Contig135 (74577-75569 p)	70	<p>Identities = 208/323 (64%), Positives = 249/323 (76%), Gaps = 1/323 (0%) gb AAK05757.1 AE006396_8 (AE006396) phosphate acetyltransferase (EC 2.3.1.8) [Lactococcus lactis subsp. lactis] Length = 326</p>
SeqID 1512	SA-2513.2	Contig138 (26798-27427 m)	72	<p>Identities = 122/204 (59%), Positives = 156/204 (75%) pir B69878 guanylate kinase homolog yloD - Bacillus subtilis emb CAA74271.1 (Y13937) putative Gmk protein [Bacillus subtilis] emb CAB13441.1 (Z99112) similar to guanylate kinase [Bacillus subtilis] Length = 244</p>

SeqID 1513	SA-2514.1	Contig138 (27594-28298 m)	33	Identities = 46/183 (25%), Positives = 81/183 (44%), Gaps = 11/183 (6%) pir F69065 hypothetical protein MTH1490 - Methanobacterium thermoautotrophicum (strain Delta H) gb AAB85965.1 (AE000909) unknown [Methanobacterium thermoautotrophicum] Length = 188
SeqID 1514	SA-2515.2	Contig138 (28348-28938 m)	78	Identities = 121/199 (60%), Positives = 157/199 (78%), Gaps = 3/199 (1%) emb CAB54585.1 (AJ006400) response regulator [Streptococcus pneumoniae] Length = 199
SeqID 1515	SA-2516.2	Contig68 (1532-2206 p)	58	Identities = 109/221 (49%), Positives = 150/221 (67%), Gaps = 9/221 (4%) gb AAK02817.1 (AE006110) unknown [Pasteurella multocida] Length = 220
SeqID 1516	SA-2517.1	Contig68 (2294-2686 p)	62	Identities = 59/131 (45%), Positives = 86/131 (65%), Gaps = 2/131 (1%) sp P44638 LGUL_HAEIN LACTOYLGLUTATHIONE LYASE (METHYLGLYOXALASE) (ALDOKETOMUTASE) (GLYOXALASE I) (GLX I) (KETONE-ALDEHYDE MUTASE) (S-D-LACTOYLGLUTATHIONE METHYLGLYOXAL LYASE) pir J64147 lactoylglutathione lyase (EC 4.4.1.5) - Haemophilus influenzae gb AAC21986.1 (U32717) lactoylglutathione lyase (gloA) [Haemophilus influenzae Rd] Length = 135
SeqID 1517	SA-2518.1	Contig68 (2815-3741 p)	64	Identities = 139/308 (45%), Positives = 202/308 (65%), Gaps = 3/308 (0%) sp Q45539 CSBB_BACSU CSBB PROTEIN pir JC5173 stress response protein csbB - Bacillus subtilis gb AAB38429.1 (L77099) 44 identity over 302 residues with hypothetical protein from Synechocystis sp. accession D64006_CD; expression induced by environmental stress; some similarity to glycosyl transferases; two potential membrane-spanning helices [Bacillus subtilis] emb CAB12688.1 (Z99108) stress response protein [Bacillus subtilis] dbj BAA24480.1 (D85082) YfhN [Bacillus subtilis] Length = 329

SeqID 1518	SA-2519.1	Contig68 (3870-4025 p)	58	<p>Identities = 30/46 (65%), Positives = 37/46 (80%) sp P45495 YPEV_LACDL HYPOTHETICAL TRANSPORT PROTEIN IN PEPV 3 REGION (ORF2) pir S57903 probable amino acid permease - Lactobacillus delbrueckii (fragment) emb CAA83253.1 (Z31377) potential amino acid permease [Lactobacillus delbrueckii] Length = 175</p>
SeqID 1519	SA-252.1	Contig138 (47259-48047 m)	50	<p>Identities = 90/224 (40%), Positives = 137/224 (60%), Gaps = 10/224 (4%) sp P54953 YXEN_BACSU PROBABLE AMINO- ACID ABC TRANSPORTER PERMEASE PROTEIN YXEN pir H70075 amino acid ABC transporter (permease) homolog yxen - Bacillus subtilis dbj BAA08330.1 (D45912) homologous to Gln transport system permease proteins [Bacillus subtilis] emb CAB15985.1 (Z99124) similar to amino acid ABC transporter (permease) [Bacillus subtilis] Length = 224</p>
SeqID 1520	SA-2520.2	Contig115 (6665-8494 p)	13	<p>Identities = 43/206 (20%), Positives = 91/206 (43%), Gaps = 12/206 (5%) pir H64496 hypothetical protein MJ1577 - Methanococcus jannaschii gb AAB99606.1 (U67598) M. jannaschii predicted coding region MJ1577 [Methanococcus jannaschii] Length = 598</p>
SeqID 1521	SA-2522.2	Contig69 (1649-1921 m)	78	<p>Identities = 81/90 (90%), Positives = 85/90 (94%) gb AAK05667.1 AE006387_10 (AE006387) 30S ribosomal protein S16 [Lactococcus lactis subsp. lactis] Length = 90</p>
SeqID 1522	SA-2523.2	Contig69 (1397-1639 m)	75	<p>Identities = 39/75 (52%), Positives = 63/75 (84%) gb AAK05666.1 AE006387_9 (AE006387) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 79</p>
SeqID 1523	SA-2524.1	Contig69 (3-1265 m)	46	<p>Identities = 115/437 (26%), Positives = 201/437 (45%), Gaps = 26/437 (5%) gb AAC97147.1 (U49397) Nra [Streptococcus pyogenes] Length = 511</p>

SeqID 1524	SA-2525.1	Contig60 (87-1895 m)	89	Identities = 463/603 (76%), Positives = 540/603 (88%) sp P37949 LEPA_BACSU GTP-BINDING PROTEIN LEPA pir G69649 GTP-binding protein lepA - Bacillus subtilis emb CAA62842.1 (X91655) lepA [Bacillus subtilis] dbj BAA12460.1 (D84432) YqeQ [Bacillus subtilis] emb CAB14493.1 (Z99117) GTP-binding protein [Bacillus subtilis] Length = 612
SeqID 1525	SA-2526.1	Contig61 (969-2801 m)	29	Identities = 95/382 (24%), Positives = 180/382 (46%), Gaps = 26/382 (6%) dbj BAB06137.1 (AP001515) DNA polymerase III (alpha subunit) [Bacillus halodurans] Length = 1433
SeqID 1526	SA-2528.1	Contig61 (82-843 p)	80	Identities = 217/232 (93%), Positives = 226/232 (96%) emb CAB90834.1 (AJ250837) putative transposase [Streptococcus dysgalactiae] Length = 259
SeqID 1527	SA-253.1	Contig138 (48157-48969 m)	49	Identities = 80/273 (29%), Positives = 136/273 (49%), Gaps = 27/273 (9%) emb CAA68052.1 (X99716) collagen binding protein [Lactobacillus reuteri] Length = 263
SeqID 1528	SA-2531.2	Contig136 (1-1152 p)	26	Identities = 51/228 (22%), Positives = 100/228 (43%), Gaps = 36/228 (15%) emb CAB39029.1 (AL034559) hypothetical protein, PFC0905c [Plasmodium falciparum] Length = 3085
SeqID 1529	SA-2532.1	Contig137 (4715-4921 p)	No Hits found	
SeqID 1530	SA-2533.1	Contig139 (122098-122286 p)	41	Identities = 18/34 (52%), Positives = 26/34 (75%) gb AAC98435.1 (L29324) unknown [Streptococcus pneumoniae] Length = 121
SeqID 1531	SA-2534.1	Contig139 (122022-122204 p)	63	Identities = 32/53 (60%), Positives = 39/53 (73%) gb AAC98435.1 (L29324) unknown [Streptococcus pneumoniae] Length = 121
SeqID 1532	SA-2535.1	Contig139 (51975-52250 p)	88	Identities = 85/91 (93%), Positives = 88/91 (96%) gb AAD40808.1 L38946_1 (L38946) histone-like DNA-binding protein [Streptococcus pyogenes] Length = 91

SeqID 1533	SA-2536.1	Contig139 (45692-46564 p)	58	Identities = 126/258 (48%), Positives = 175/258 (66%), Gaps = 2/258 (0%) sp O66126 SPA_MICLU GERANYLTRANSFERASE (FARNESYL-DIPHOSPHATE SYNTHASE) (FPP SYNTHASE) dbj BAA25265.1 (AB003187) farnesyl diphosphate synthase [Micrococcus luteus] Length = 291
SeqID 1534	SA-2537.1	Contig139 (38436-38666 p)	39	Identities = 22/77 (28%), Positives = 41/77 (52%), Gaps = 4/77 (5%) pir G69865 hypothetical protein ykuJ - Bacillus subtilis emb CAA10873.1 (AJ222587) YkuJ protein [Bacillus subtilis] emb CAB13283.1 (Z99111) ykuJ [Bacillus subtilis] Length = 79
SeqID 1535	SA-2539.1	Contig131 (7218-7448 p)	No Hits found	
SeqID 1536	SA-2540.2	Contig138 (95981-96244 p)	No Hits found	
SeqID 1537	SA-2542.1	Contig133 (27162-27320 m)	No Hits found	
SeqID 1538	SA-2543.1	Contig132 (25696-25839 m)	No Hits found	
SeqID 1539	SA-2545.1	Contig127 (19934-20134 p)	No Hits found	
SeqID 1540	SA-2547.1	Contig135 (30113-30319 p)	No Hits found	
SeqID 1541	SA-2548.1	Contig135 (29845-30003 p)	No Hits found	
SeqID 1542	SA-2549.1	Contig135 (27910-28200 m)	No Hits found	
SeqID 1543	SA-255.1	Contig138 (49107-50450 m)	67	Identities = 215/439 (48%), Positives = 311/439 (69%), Gaps = 5/439 (1%) sp P54475 YQFR_BACSU PROBABLE RNA HELICASE IN CCCA-SODA INTERGENIC REGION pir D69954 ATP-dependent RNA helicase homolog yqfR - Bacillus subtilis dbj BAA12495.1 (D84432) YqfR [Bacillus subtilis] emb CAB14444.1 (Z99116) similar to ATP-dependent RNA helicase [Bacillus subtilis] Length = 438
SeqID 1544	SA-2551.2	Contig123 (36802-36984 m)	No Hits found	
SeqID 1545	SA-2554.1	Contig116 (30459-30827 m)	No Hits found	
SeqID 1546	SA-2555.1	Contig116 (60-398 m)	No Hits found	
SeqID 1547	SA-2556.1	Contig134 (78113-78304 p)	45	Identities = 24/75 (32%), Positives = 46/75 (61%) sp O32233 SECG_BACSU PROBABLE PROTEIN-EXPORT MEMBRANE PROTEIN SECG pir A70028 hypothetical protein yvaL - Bacillus subtilis emb CAB15368.1 (Z99121) yvaL [Bacillus subtilis] Length = 76

SeqID 1548	SA-2557.1	Contig134 (69629-69823 m)	No Hits found	
SeqID 1549	SA-2558.1	Contig134 (69416-69646 p)	No Hits found	
SeqID 1550	SA-2559.1	Contig139 (132392-132580 m)	61	Identities = 36/46 (78%), Positives = 39/46 (84%) gb AAC38687.1 (AF030361) transposase [Streptococcus pneumoniae] gb AAC38702.1 (AF030364) transposase [Streptococcus pneumoniae] Length = 418
SeqID 1551	SA-256.1	Contig138 (50548-51558 m)	79	Identities = 221/336 (65%), Positives = 271/336 (79%), Gaps = 11/336 (3%) sp Q9ZHA5 MRAY_STRPN PHOSPHO-N-ACETYL-MURAMOYL-PENTAPEPTIDE-TRANSFERASE (UDP-MURNAC-PENTAPEPTIDE PHOSPHOTRANSFERASE) gb AAC95457.1 (AF068903) undecaprenyl-phosphate-UDP-MurNac-pentapeptide phospho-MurNac-pentapeptide transferase [Streptococcus pneumoniae] Length = 326
SeqID 1552	SA-2560.1	Contig139 (133610-133768 p)	37	Identities = 23/35 (65%), Positives = 28/35 (79%) gb AAG38044.1 AF295925_9 (AF295925) Orf28 [Streptococcus pneumoniae] Length = 371
SeqID 1553	SA-2561.1	Contig139 (154195-154389 p)	No Hits found	
SeqID 1554	SA-2562.1	Contig125 (24635-24835 p)	No Hits found	
SeqID 1555	SA-2564.1	Contig112 (19746-19868 p)	47	Identities = 22/40 (55%), Positives = 27/40 (67%), Gaps = 1/40 (2%) dbj BAB04083.1 (AP001508) transcriptional regulator [Bacillus halodurans] Length = 66
SeqID 1556	SA-2565.1	Contig112 (19523-19741 p)	No Hits found	
SeqID 1557	SA-2566.1	Contig111 (15881-16075 p)	No Hits found	
SeqID 1558	SA-2567.1	Contig111 (13119-13307 p)	No Hits found	
SeqID 1559	SA-257.1	Contig138 (51560-53821 m)	73	Identities = 405/741 (54%), Positives = 556/741 (74%), Gaps = 10/741 (1%) gb AAAF17266.1 AF210756_1 (AF210756) penicillin-binding protein 2X [Streptococcus pneumoniae] Length = 750
SeqID 1560	SA-2571.1	Contig106 (21708-21950 p)	41	Identities = 32/83 (38%), Positives = 46/83 (54%) sp P55661 Y4TG_RHISN PROBABLE AMINO-ACID ABC TRANSPORTER PERMEASE PROTEIN Y4TG gb AAB91860.1 (AE000098) Y4tG [Rhizobium sp. NGR234] Length = 231

SeqID 1561	SA-2572.1	Contig106 (21244-21573 p)	79	Identities = 79/110 (71%), Positives = 91/110 (81%), Gaps = 1/110 (0%) sp P16680 PHNA_ECOLI PHNA PROTEIN pir B35718 alkylphosphonate uptake protein phnA [imported] - Escherichia coli gb AAA24337.1 (J05260) phnA protein [Escherichia coli] gb AAC97007.1 (U14003) phnA gene product [Escherichia coli] gb AAC77069.1 (AE000483) orf, hypothetical protein [Escherichia coli K12] Length = 111
SeqID 1562	SA-2574.1	Contig135 (83485-83640 p)	No Hits found	
SeqID 1563	SA-2575.1	Contig125 (16287-16571 m)	No Hits found	
SeqID 1564	SA-2576.2	Contig134 (50206-50895 p)	85	Identities = 190/228 (83%), Positives = 204/228 (89%) gb AAK06098.1 AE006430_3 (AE006430) 50S ribosomal protein L1 [Lactococcus lactis subsp. lactis] Length = 229
SeqID 1565	SA-2577.1	Contig135 (69410-69703 m)	42	Identities = 35/72 (48%), Positives = 42/72 (57%) pir S59084 hypothetical protein 29.1 - red alga (Chondrus crispus) mitochondrion emb CAA87600.1 (Z47547) unique orf [Chondrus crispus] Length = 79
SeqID 1566	SA-258.1	Contig138 (53822-54148 m)	50	Identities = 44/99 (44%), Positives = 71/99 (71%) emb CAB01928.1 (Z79691) FtsL [Streptococcus pneumoniae] gb AAC95455.1 (AF068903) YIID [Streptococcus pneumoniae] Length = 105
SeqID 1567	SA-2582.1	Contig89 (20-172 m)	53	Identities = 27/42 (64%), Positives = 35/42 (83%) gb AAD09220.1 (U74080) unknown [Streptococcus gordonii] Length = 50
SeqID 1568	SA-2583.1	Contig133 (74342-74776 m)	62	Identities = 51/145 (35%), Positives = 91/145 (62%), Gaps = 3/145 (2%) pir H70091 hypothetical protein yydJ - Bacillus subtilis dbj BAA11272.1 (D78193) yydJ [Bacillus subtilis] emb CAB16051.1 (Z99124) yydJ [Bacillus subtilis] Length = 240
SeqID 1569	SA-2584.1	Contig133 (74719-75069 m)	60	Identities = 47/106 (44%), Positives = 71/106 (66%) pir H70091 hypothetical protein yydJ - Bacillus subtilis dbj BAA11272.1 (D78193) yydJ [Bacillus subtilis] emb CAB16051.1 (Z99124) yydJ [Bacillus subtilis] Length = 240

SeqID 1570	SA-2586.1	Contig80 (961-1092 m)	No Hits found	
SeqID 1571	SA-2588.2	Contig96 (1977-2162 p)	No Hits found	
SeqID 1572	SA-259.1	Contig138 (54163-55110 m)	84	Identities = 265/316 (83%), Positives = 288/316 (90%) gb AAC95454.1 (AF068903) Y1C [Streptococcus pneumoniae] Length = 316
SeqID 1573	SA-2593.1	Contig65 (2145-2339 p)	No Hits found	
SeqID 1574	SA-2594.1	Contig62 (165-344 p)	No Hits found	
SeqID 1575	SA-2596.1	Contig91 (939-1604 p)	75	Identities = 117/217 (53%), Positives = 168/217 (76%) sp P42399 YCKA_BACSU PROBABLE AMINO-ACID ABC TRANSPORTER PERMEASE PROTEIN YCKA pir C69760 amino acid ABC transporter (permease) homolog yckA - Bacillus subtilis dbj BAA06425.1 (D30762) homolog of glutamine permease of H. influenzae [Bacillus subtilis] dbj BAA08971.1 (D50453) homolog of glutamine permease of H. influenzae [Bacillus subtilis] emb CAB12131.1 (Z99105) similar to amino acid ABC transporter (permease) [Bacillus subtilis] Length = 226
SeqID 1576	SA-2597.1	Contig91 (58-915 p)	62	Identities = 127/276 (46%), Positives = 183/276 (66%), Gaps = 12/276 (4%) sp P42400 YCKB_BACSU PROBABLE ABC TRANSPORTER EXTRACELLULAR BINDING PROTEIN YCKB PRECURSOR (ORF2) pir D69760 amino acid ABC transporter (binding protein) homolog yckB - Bacillus subtilis emb CAB12132.1 (Z99105) similar to amino acid ABC transporter (binding protein) [Bacillus subtilis] Length = 287

SeqID 1577	SA-2598.1	Contig104 (8256-8870 p)	67	Identities = 109/218 (50%), Positives = 146/218 (66%), Gaps = 15/218 (6%) sp P46338 YQGG_BACSU PROBABLE ABC TRANSPORTER BINDING PROTEIN YQGG PRECURSOR pir A69956 phosphate ABC transporter (binding protein) homolog yqgG - Bacillus subtilis dbj BAA09581.1 (D58414) ORF108 [Bacillus subtilis] dbj BAA12510.1 (D84432) YqgG [Bacillus subtilis] emb CAB14429.1 (Z99116) alternate gene name: yzmB~similar to phosphate ABC transporter (binding protein) [Bacillus subtilis] Length = 300
SeqID 1578	SA-2599.1	Contig135 (79561-80901 p)	54	Identities = 153/445 (34%), Positives = 250/445 (55%), Gaps = 11/445 (2%) emb CAB61253.1 (AJ250422) ORFC [Oenococcus oeni] Length = 463
SeqID 1579	SA-26.1	Contig137 (15639-15896 p)	No Hits found	Identities = 303/417 (72%), Positives = 356/417 (84%), Gaps = 1/417 (0%) sp P96489 PROA_STRTR GAMMA-GLUTAMYL PHOSPHATE REDUCTASE (GPR) (GLUTAMATE-5-SEMIALDEHYDE DEHYDROGENASE) (GLUTAMYL-GAMMA-SEMIALDEHYDE DEHYDROGENASE) emb CAA63148.1 (X92418) gamma-glutamyl phosphate reductase [Streptococcus thermophilus] Length = 416
SeqID 1580	SA-260.1	Contig138 (55193-56446 m)	84	
SeqID 1581	SA-2600.1	Contig110 (3263-3451 m)	No Hits found	
SeqID 1582	SA-2602.1	Contig110 (3909-4259 p)	70	Identities = 63/115 (54%), Positives = 83/115 (71%), Gaps = 1/115 (0%) sp O35046 YOC_D_BACSU HYPOTHETICAL 36.4 KDA PROTEIN IN CSA-DES INTERGENIC REGION pir A69901 immunity to bacteriotoxins homolog yocD - Bacillus subtilis gb AAB84435.1 (AF027868) YocD [Bacillus subtilis] emb CAB13809.1 (Z99114) similar to immunity to bacteriotoxins [Bacillus subtilis] Length = 325
SeqID 1583	SA-2603.1	Contig111 (11745-11921 p)	No Hits found	

SeqID 1584	SA-2604.1	Contig116 (30551-30826 m)	34	Identities = 30/51 (58%), Positives = 32/51 (61%) pir F71245 hypothetical protein PHS004 - Pyrococcus horikoshii dbj BAA29293.1 (AP000001) 58aa long hypothetical protein [Pyrococcus horikoshii] Length = 58
SeqID 1585	SA-2605.1	Contig116 (7428-7886 p)	80	Identities = 122/123 (99%), Positives = 123/123 (99%) gb AAG09968.1 AF248037_3 (AF248037) cation efflux system protein [Streptococcus agalactiae] Length = 287
SeqID 1586	SA-2606.1	Contig138 (96371-96607 p)	No Hits found	
SeqID 1587	SA-2608.1	Contig139 (135435-135725 m)	No Hits found	
SeqID 1588	SA-2609.1	Contig139 (132781-132939 m)	No Hits found	
SeqID 1589	SA-261.1	Contig138 (56456-57259 m)	86	Identities = 200/265 (75%), Positives = 235/265 (88%) sp P96488 PROB_STRTR GLUTAMATE 5-KINASE (GAMMA-GLUTAMYL KINASE) (GK) emb CAA63147.1 (X92418) gamma-glutamyl kinase [Streptococcus thermophilus] Length = 267
SeqID 1590	SA-2610.1	Contig139 (132368-132517 p)	60	Identities = 19/36 (52%), Positives = 30/36 (82%) pir T50042 hypothetical protein tasA [imported] - Streptococcus pneumoniae emb CAA59773.1 (X85787) tasA [Streptococcus pneumoniae] Length = 359
SeqID 1591	SA-2612.1	Contig139 (93569-93832 p)	No Hits found	
SeqID 1592	SA-2615.1	Contig139 (52350-52646 m)	No Hits found	
SeqID 1593	SA-2618.1	Contig139 (18458-18772 m)	35	Identities = 30/71 (42%), Positives = 37/71 (51%), Gaps = 1/71 (1%) pir G72514 hypothetical protein APE2092 - Aeropyrum pernix (strain K1) dbj BAA81103.1 (AP000063) 101aa long hypothetical protein [Aeropyrum pernix] Length = 101
SeqID 1594	SA-2620.1	Contig138 (81930-82160 m)	No Hits found	

SeqID 1595	SA-263.1	Contig138 (57400-58785 p)	57	Identities = 174/447 (38%), Positives = 267/447 (58%), Gaps = 10/447 (2%) pir T47097 hypothetical protein N17C [imported] - Bacillus subtilis dbj BAA06652.1 (D31856) hypothetical protein [Bacillus subtilis] dbj BAA06256.1 (D29985) hypothetical 64.7-kDa protein [Bacillus subtilis] emb CAB15963.1 (Z99124) phosphotransferase system (PTS) beta-glucoside-specific enzyme IIABC component [Bacillus subtilis] Length = 609
SeqID 1596	SA-2632.1	Contig137 (2699-2956 p)	No Hits found	
SeqID 1597	SA-2635.1	Contig136 (45825-45980 p)	No Hits found	
SeqID 1598	SA-264.1	Contig138 (58993-60597 m)	No Hits found	
SeqID 1599	SA-2642.1	Contig135 (29120-29389 m)	No Hits found	
SeqID 1600	SA-2646.1	Contig134 (73407-73562 m)	No Hits found	
SeqID 1601	SA-2649.1	Contig133 (70298-70522 p)	No Hits found	
SeqID 1602	SA-265.1	Contig138 (60601-61335 m)	65	Identities = 95/243 (39%), Positives = 164/243 (67%), Gaps = 2/243 (0%) pir H75077 abc transporter, ATP-binding protein PAB1696 - Pyrococcus abyssi (strain Orsay) emb CAB49925.1 (AJ248286) ABC transporter, ATP-binding protein [Pyrococcus abyssi] Length = 253
SeqID 1603	SA-2650.1	Contig133 (68964-69245 p)	No Hits found	
SeqID 1604	SA-2651.1	Contig133 (60101-60367 p)	No Hits found	
SeqID 1605	SA-2654.1	Contig132 (48100-48423 m)	34	Identities = 34/73 (46%), Positives = 37/73 (50%), Gaps = 3/73 (4%) pir E71186 hypothetical protein PH1769 - Pyrococcus horikoshii dbj BAA30884.1 (AP000007) 100aa long hypothetical protein [Pyrococcus horikoshii] Length = 100
SeqID 1606	SA-2655.1	Contig132 (48095-48463 p)	93	Identities = 116/122 (95%), Positives = 120/122 (98%) sp Q9WVW2 RL14_STRPN 50S RIBOSOMAL PROTEIN L14 gb AAD33266.1 AF126059_7 (AF126059) Rpl14 [Streptococcus pneumoniae] gb AAD33275.1 (AF126060) Rpl14 [Streptococcus pneumoniae] gb AAD33284.1 (AF126061) Rpl14 [Streptococcus pneumoniae] Length = 122

SeqID 1607	SA-2657.1	Contig132 (25921-26223 p)	No Hits found	
SeqID 1608	SA-266.1	Contig138 (61355-61660 m)	No Hits found	
SeqID 1609	SA-2664.1	Contig131 (2878-3129 m)	No Hits found	
SeqID 1610	SA-2666.1	Contig130 (70425-70709 p)	No Hits found	
SeqID 1611	SA-2668.1	Contig130 (38146-38331 m)	No Hits found	
SeqID 1612	SA-2669.1	Contig130 (27964-28215 p)	No Hits found	
SeqID 1613	SA-267.1	Contig138 (61873-63858 m)	77	Identities = 403/661 (60%), Positives = 520/661 (77%), Gaps = 8/661 (1%) dbj BAB06071.1 (AP001515) transketolase [Bacillus halodurans] Length = 666
SeqID 1614	SA-2670.1	Contig129 (57906-58148 m)	No Hits found	
SeqID 1615	SA-2672.1	Contig129 (47951-48133 p)	No Hits found	
SeqID 1616	SA-2673.1	Contig129 (21896-22162 p)	53	Identities = 40/83 (48%), Positives = 64/83 (76%) pir A69742 conserved hypothetical protein yzaA - Bacillus subtilis emb CAB11811.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 99
SeqID 1617	SA-268.1	Contig138 (63983-65413 m)	18	Identities = 48/191 (25%), Positives = 88/191 (45%), Gaps = 9/191 (4%) emb CAB65412.1 (AJ243530) putative Mga-like regulatory protein [Streptococcus dysgalactiae subsp. dysgalactiae] Length = 497
SeqID 1618	SA-2681.1	Contig126 (41736-42008 p)	No Hits found	
SeqID 1619	SA-2683.1	Contig126 (33868-34071 p)	No Hits found	
SeqID 1620	SA-2687.1	Contig125 (36869-37171 m)	No Hits found	
SeqID 1621	SA-2689.1	Contig124 (23609-23926 p)	33	Identities = 25/72 (34%), Positives = 36/72 (49%) pir G72510 hypothetical protein APE2061 - Aeropyrum pernix (strain K1) dbj BAA81071.1 (AP000063) 114aa long hypothetical protein [Aeropyrum pernix] Length = 114
SeqID 1622	SA-269.1	Contig138 (65404-66741 m)	44	Identities = 105/423 (24%), Positives = 198/423 (45%), Gaps = 15/423 (3%) sp P37061 NAOX_ENTFA NADH OXIDASE (NOXASE) pir S26965 NADH oxidase - Enterococcus faecalis emb CAA48728.1 (X68847) NADH oxidase [Enterococcus faecalis] Length = 446
SeqID 1623	SA-2690.1	Contig123 (39402-39596 p)	No Hits found	

SeqID 1624	SA-2691.1	Contig123 (39313-39621 p)	36	Identities = 21/61 (34%), Positives = 38/61 (61%), Gaps = 1/61 (1%) ref NP_076769.1 Orf18 [bacteriophage bIL310] gb AAK04132.1 AE006242_11 (AE006242) prophage ps1 protein 10 [Lactococcus lactis subsp. lactis] gb AAK08422.1 AF323671_18 (AF323671) Orf18 [bacteriophage bIL310] Length = 93
SeqID 1625	SA-2696.1	Contig122 (25328-25630 m)	34	Identities = 22/56 (39%), Positives = 35/56 (62%), Gaps = 4/56 (7%) sp P36417 GBF_DICDI G-BOX BINDING FACTOR (GBF) pir J53185 G-box-binding factor - slime mold (Dictyostelium discoideum) gb AAA21021.1 (L29075) G-box binding factor [Dictyostelium discoideum] Length = 708
SeqID 1626	SA-270.1	Contig138 (66829-67542 m)	81	Identities = 150/230 (65%), Positives = 194/230 (84%), Gaps = 1/230 (0%) sp P52281 GLPF_STRPN GLYCEROL UPTAKE FACILITATOR PROTEIN pir S67937 glycerol uptake facilitator GlpF - Streptococcus pneumoniae (strain P13) gb AA91618.1 (U12567) glycerol uptake facilitator [Streptococcus pneumoniae] Length = 233
SeqID 1627	SA-2703.1	Contig119 (308-1129 m)	50	Identities = 94/304 (30%), Positives = 152/304 (49%), Gaps = 32/304 (10%) gb AAK04228.1 AE006251_4 (AE006251) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 304
SeqID 1628	SA-2705.1	Contig118 (27339-27527 p)	No Hits found	
SeqID 1629	SA-2707.1	Contig118 (17969-18169 p)	80	Identities = 52/66 (78%), Positives = 60/66 (90%) gb AAK05941.1 AE006414_7 (AE006414) 50S ribosomal protein L35 [Lactococcus lactis subsp. lactis] Length = 66
SeqID 1630	SA-2709.1	Contig118 (17399-17929 p)	74	Identities = 112/169 (66%), Positives = 134/169 (79%) sp O53084 IF3_LISMO TRANSLATION INITIATION FACTOR IF- 3 emb CAA68920.1 (Y07640) translation initiation factor, IF3 [Listeria monocytogenes] Length = 171
SeqID 1631	SA-271.1	Contig138 (67539-69368 m)	87	Identities = 464/608 (76%), Positives = 539/608 (88%) gb AAC34740.1 (U94770) alpha-glycerophosphate oxidase [Streptococcus pneumoniae] Length = 608

SeqID 1632	SA-2710.1	Contig118 (16555-17238 p)	73	Identities = 139/223 (62%), Positives = 171/223 (76%), Gaps = 3/223 (1%) gb AAK05801.1 AE006401_2 (AE006401) cytidine monophosphate kinase (EC 2.7.4.14) [Lactococcus lactis subsp. lactis] Length = 220
SeqID 1633	SA-2711.1	Contig118 (16244-16408 m)	No Hits found	
SeqID 1634	SA-2712.1	Contig118 (16020-16544 p)	42	Identities = 46/167 (27%), Positives = 82/167 (48%), Gaps = 15/167 (8%) gb AAK05013.1 AE006326_4 (AE006326) teichoic acid ABC transporter ATP binding protein [Lactococcus lactis subsp. lactis] Length = 466
SeqID 1635	SA-2715.1	Contig117 (11061-11336 p)	No Hits found	
SeqID 1636	SA-2718.1	Contig116 (21405-21707 p)	No Hits found	
SeqID 1637	SA-2719.1	Contig116 (19624-19911 p)	No Hits found	
SeqID 1638	SA-272.1	Contig138 (69381-70889 m)	85	Identities = 383/501 (76%), Positives = 442/501 (87%), Gaps = 1/501 (0%) sp O34154 GLPK_ENTFA GLYCEROL KINASE (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE) (GLYCEROKINASE) (GK) gb AAB69986.1 (U94356) glycerol kinase [Enterococcus faecalis] Length = 501
SeqID 1639	SA-2722.1	Contig115 (26094-26351 m)	45	Identities = 36/58 (62%), Positives = 39/58 (67%) gb AAG12204.1 AF287482_5 (AF287482) Orf122 [Chlorobium tepidum] Length = 121
SeqID 1640	SA-2726.1	Contig113 (25586-25693 p)	No Hits found	
SeqID 1641	SA-273.1	Contig138 (71003-71266 m)	43	Identities = 36/79 (45%), Positives = 48/79 (60%), Gaps = 1/79 (1%) pir G83401 hypothetical protein PA1960 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG05348.1 AE004622_10 (AE004622) hypothetical protein [Pseudomonas aeruginosa] Length = 240
SeqID 1642	SA-2733.1	Contig108 (18783-18911 p)	No Hits found	
SeqID 1643	SA-274.1	Contig138 (71355-71612 m)	58	Identities = 41/72 (56%), Positives = 56/72 (76%) pir E69894 hypothetical protein ynzC - Bacillus subtilis emb CAB13672.1 (Z99113) ynzC [Bacillus subtilis] Length = 77
SeqID 1644	SA-2741.1	Contig102 (13042-13308 p)	No Hits found	
SeqID 1645	SA-2747.1	Contig98 (8006-8278 p)	No Hits found	

SeqID 1646	SA-275.1	Contig138 (71624-73663 m)	65	Identities = 315/687 (45%), Positives = 445/687 (63%), Gaps = 21/687 (3%) sp P54381 SYGB_BACSU GLYCYL-TRNA SYNTHETASE BETA CHAIN (GLYCINE--TRNA LIGASE BETA CHAIN) (GLYRS) pir B69636 glycine--tRNA ligase (EC 6.1.1.14) beta chain glyS - Bacillus subtilis dbj BAA12485.1 (D84432) YqfK [Bacillus subtilis] emb CAB14455.1 (Z99116) glycyl-tRNA synthetase (beta subunit) [Bacillus subtilis] emb CAB14468.1 (Z99117) glycyl-tRNA synthetase (beta subunit) [Bacillus subtilis] Length = 679
SeqID 1647	SA-2752.1	Contig94 (10312-10587 p)	No Hits found	Identities = 32/79 (40%), Positives = 48/79 (60%) gb AAK05381.1 AE006360_10 (AE006360) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 79
SeqID 1648	SA-2756.1	Contig91 (6477-6716 p)	60	
SeqID 1649	SA-2757.1	Contig91 (6359-6622 m)	No Hits found	
SeqID 1650	SA-2758.1	Contig91 (5806-6126 m)	No Hits found	
SeqID 1651	SA-2759.1	Contig91 (5708-6319 p)	46	Identities = 66/153 (43%), Positives = 94/153 (61%), Gaps = 2/153 (1%) gb AAK05259.1 AE006348_4 (AE006348) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 199
SeqID 1652	SA-276.1	Contig138 (73667-74164 m)	45	Identities = 64/176 (36%), Positives = 99/176 (55%), Gaps = 6/176 (3%) gb AAD24436.1 AF112858_1 (AF112858) NAD(P)H dehydrogenase [Bacillus stearothermophilus] Length = 211
SeqID 1653	SA-2760.1	Contig91 (613-960 m)	45	Identities = 33/106 (31%), Positives = 53/106 (49%), Gaps = 8/106 (7%) emb CAA33190.1 (X15081) MURF2 protein (AA 1-348) [Crithidia fasciculata] Length = 347
SeqID 1654	SA-2761.1	Contig90 (10903-11157 m)	No Hits found	
SeqID 1655	SA-2763.1	Contig90 (3-899 p)	99	Identities = 298/298 (100%), Positives = 298/298 (100%) gb AAK14387.1 AF338416_1 (AF338416) glyceraldehyde 3-phosphate dehydrogenase [Streptococcus agalactiae] Length = 336
SeqID 1656	SA-2766.1	Contig85 (11550-11849 m)	No Hits found	

SeqID 1657	SA-2767.1	Contig84 (10458-10913 m)	67	Identities = 79/142 (55%), Positives = 102/142 (71%), Gaps = 5/142 (3%) gb AAK05046.1 AE006329_5 (AE006329) UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis] Length = 159
SeqID 1658	SA-2768.1	Contig84 (10038-10448 m)	50	Identities = 38/131 (29%), Positives = 70/131 (53%), Gaps = 5/131 (3%) pir B83475 hypothetical protein PA1353 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG04742.1 AE004565_1 (AE004565) hypothetical protein [Pseudomonas aeruginosa] Length = 137
SeqID 1659	SA-2769.1	Contig83 (5598-5870 p)	No Hits found	
SeqID 1660	SA-277.1	Contig138 (74308-75225 m)	72	Identities = 221/287 (77%), Positives = 250/287 (87%) sp Q9KDD9 SYGA_BACHD GLYCYL-TRNA SYNTHETASE ALPHA CHAIN (GLYCINE--TRNA LIGASE ALPHA CHAIN) (GLYRS) dbj BAB05089.1 (AP001511) glycyL-TRNA synthetase (alpha subunit) [Bacillus halodurans] Length = 297
SeqID 1661	SA-2770.1	Contig83 (3561-3887 p)	24	Identities = 17/58 (29%), Positives = 27/58 (46%), Gaps = 1/58 (1%) ref NP_042981.1 U88 [Human herpesvirus 6] emb CAA58337.1 (X83413) U88 [Human herpesvirus 6] Length = 413
SeqID 1662	SA-2771.1	Contig80 (1912-2166 m)	No Hits found	
SeqID 1663	SA-2772.1	Contig80 (902-1168 p)	No Hits found	
SeqID 1664	SA-2773.1	Contig78 (6209-6433 p)	49	Identities = 26/51 (50%), Positives = 37/51 (71%), Gaps = 2/51 (3%) gb AAK04732.1 AE006296_6 (AE006296) DNA-directed RNA polymerase delta chain (EC 2.7.7.6) [Lactococcus lactis subsp. lactis] Length = 187
SeqID 1665	SA-2776.1	Contig69 (2000-2224 p)	No Hits found	
SeqID 1666	SA-278.2	Contig138 (75535-76077 m)	57	Identities = 71/168 (42%), Positives = 105/168 (62%) pir F81147 probable integral membrane protein NMA1102 [imported] - Neisseria meningitidis (group B strain MD58, group A strain Z2491) gb AAF41294.1 (AE002440) conserved hypothetical protein [Neisseria meningitidis MC58] emb CAB84365.1 (AL162755) putative integral membrane protein [Neisseria meningitidis Z2491] Length = 169

SeqID 1667	SA-279.1	Contig126 (42614-42796 p)	No Hits found	
SeqID 1668	SA-28.1	Contig137 (13147-15513 p)	20	Identities = 96/354 (27%), Positives = 161/354 (45%), Gaps = 61/354 (17%) ref NP_053238.1 pXO2-84 [Bacillus anthracis] gb AAF13688.1 AF188935_86 (AF188935) pXO2-84 [Bacillus anthracis] Length = 490
SeqID 1669	SA-280.1	Contig126 (42553-46098 m)	69	Identities = 629/1241 (50%), Positives = 825/1241 (65%), Gaps = 75/1241 (6%) gb AAG33958.1 AF217414_1 (AF217414) pullulanase [Streptococcus pneumoniae] Length = 1287
SeqID 1670	SA-282.1	Contig126 (41501-42421 m)	61	Identities = 134/299 (44%), Positives = 197/299 (65%) dbj BAB08178.1 (AB036768) exfoliative toxin A [Staphylococcus hyicus] Length = 306
SeqID 1671	SA-283.1	Contig126 (41283-41459 p)	No Hits found	
SeqID 1672	SA-285.1	Contig126 (40282-41184 m)	62	Identities = 139/311 (44%), Positives = 200/311 (63%), Gaps = 21/311 (6%) sp Q9KAC3 MIAA_BACHD TRNA DELTA(2)-ISOPENTENYL PYROPHOSPHATE TRANSFERASE (IPP TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE) (IPTASE) (IPT) dbj BAB06085.1 (AP001515) tRNA isopentenylpyrophosphate transferase [Bacillus halodurans] Length = 314
SeqID 1673	SA-286.1	Contig126 (38953-40191 m)	61	Identities = 183/406 (45%), Positives = 255/406 (62%), Gaps = 12/406 (2%) dbj BAB06081.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 418
SeqID 1674	SA-287.1	Contig126 (38313-38960 m)	No Hits found	
SeqID 1675	SA-288.1	Contig126 (37342-38271 m)	64	Identities = 156/309 (50%), Positives = 209/309 (67%), Gaps = 5/309 (1%) sp P54548 YQJK_BACSU HYPOTHETICAL 34.0 KDA PROTEIN IN GLNQ-ANSR INTERGENIC REGION pir C69964 conserved hypothetical protein yqjK - Bacillus subtilis dbj BAA12617.1 (D84432) YqjK [Bacillus subtilis] emb CAB14316.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] Length = 307
SeqID 1676	SA-29.1	Contig137 (12635-13093 p)	No Hits found	

SeqID 1677	SA-290.1	Contig126 (36579-37340 m)	59	Identities = 100/253 (39%), Positives = 152/253 (59%), Gaps = 2/253 (0%) sp P54554 YQJQ_BACSU HYPOTHETICAL OXIDOREDUCTASE IN ANSR-BMRU INTERGENIC REGION pir A69965 ketoacyl reductase homolog yqjQ - Bacillus subtilis dbj BAA12623.1 (D84432) YqjQ [Bacillus subtilis] emb CAB14310.1 (Z99116) similar to ketoacyl reductase [Bacillus subtilis] Length = 259
SeqID 1678	SA-291.1	Contig126 (34384-36582 m)	60	Identities = 276/772 (35%), Positives = 447/772 (57%), Gaps = 45/772 (5%) pir H69980 single-strand DNA-specific exonuclease homolog yrvE - Bacillus subtilis emb CAB14721.1 (Z99118) similar to single-strand DNA-specific exonuclease [Bacillus subtilis] Length = 786
SeqID 1679	SA-292.1	Contig126 (31668-34187 m)	12	Identities = 65/220 (29%), Positives = 104/220 (46%), Gaps = 33/220 (15%) sp P11701 SACB_STRMU LEVANSUCRASE PRECURSOR (BETA-D-FRUCTOFURANOSYL TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE) pir B28551 levansucrase (EC 2.4.1.10) precursor - Streptococcus mutans (strain GS-5) gb AAA88584.1 (M18954) fructosyltransferase [Streptococcus mutans] Length = 797
SeqID 1680	SA-293.1	Contig126 (31027-31545 m)	74	Identities = 110/170 (64%), Positives = 135/170 (78%) sp O34443 APT_BACSU ADENINE PHOSPHORIBOSYLTRANSFERASE (APRT) pir B69587 adenine phosphoribosyltransferase (EC 2.4.2.7) - Bacillus subtilis gb AAC46040.1 (U86377) adenine phosphoribosyltransferase; Apt [Bacillus subtilis] emb CAB14720.1 (Z99118) adenine phosphoribosyltransferase [Bacillus subtilis] Length = 170
SeqID 1681	SA-294.1	Contig126 (30229-30909 m)	78	Identities = 140/227 (61%), Positives = 179/227 (78%) dbj BAA11244.1 (D78182) ORF2 [Streptococcus mutans] Length = 231
SeqID 1682	SA-295.1	Contig126 (29442-30125 m)	71	Identities = 132/226 (58%), Positives = 168/226 (73%) dbj BAA11245.1 (D78182) ORF3 [Streptococcus mutans] Length = 232

SeqID 1683	SA-296.1	Contig126 (28664-29455 m)	83	Identities = 186/262 (70%), Positives = 224/262 (84%) dbj BAA11246.1 (D78182) ORF4 [Streptococcus mutans] Length = 262
SeqID 1684	SA-297.1	Contig126 (27552-28655 m)	55	Identities = 147/368 (39%), Positives = 210/368 (56%), Gaps = 13/368 (3%) sp O32159 YURR_BACSU HYPOTHETICAL 39.4 KD OXIDOREDUCTASE IN HOM-MRGA INTERGENIC REGION pir A70019 opine catabolism homolog yurR - Bacillus subtilis emb CAB15253.1 (Z99120) similar to opine catabolism [Bacillus subtilis] Length = 372
SeqID 1685	SA-298.3	Contig126 (26639-27493 m)	89	Identities = 247/275 (89%), Positives = 260/275 (93%) gb AAD19913.1 (AF105113) glucose-1-phosphate thymidyl transferase [Streptococcus pneumoniae] Length = 289
SeqID 1686	SA-299.3	Contig126 (26032-26625 m)	86	Identities = 162/198 (81%), Positives = 183/198 (91%) gb AAC78675.1 (AF094575) dTDP-4-keto-6-deoxyglucose-3,5-epimerase Cps19aM [Streptococcus pneumoniae] Length = 198
SeqID 1687	SA-3.1	Contig137 (40498-42447 p)	50	Identities = 188/659 (28%), Positives = 327/659 (49%), Gaps = 29/659 (4%) ref NP_053164.1 pXO2.09 [Bacillus anthracis] gb AAF13614.1 AF188935_12 (AF188935) pXO2.09 [Bacillus anthracis] Length = 643
SeqID 1688	SA-30.1	Contig137 (11818-12618 p)	22	Identities = 48/115 (41%), Positives = 64/115 (54%), Gaps = 3/115 (2%) emb CAA59264.1 (X84793) streptodomainase [Streptococcus pyogenes] Length = 385
SeqID 1689	SA-301.1	Contig126 (24779-25825 m)	91	Identities = 325/347 (93%), Positives = 340/347 (97%) sp P95780 RMLB_STRMU DTDG-GLUCOSE 4,6-DEHYDRATASE dbj BAA11249.1 (D78182) dTDP-glucose-4,6-dehydratase [Streptococcus mutans] Length = 348
SeqID 1690	SA-303.1	Contig126 (22369-22569 m)	No Hits found	
SeqID 1691	SA-305.1	Contig126 (21509-24727 p)	99	Identities = 1070/1072 (99%), Positives = 1071/1072 (99%) emb CAA75865.1 (Y15903) hyaluronate lyase [Streptococcus agalactiae] Length = 1072

SeqID 1692	SA-306.1	Contig126 (20810-21292 m)	88	Identities = 132/160 (82%), Positives = 146/160 (90%), Gaps = 1/160 (0%) sp P95781 MUTX_STRMU_MUTATOR MUTT PROTEIN (7,8-DIHYDRO-8-OXOGUANINE-TRIPHOSPHATASE) (8-OXO-DGTPASE) (DGTP PYROPHOSPHOHYDROLASE) dbj BAA11250.1 (D78182) MutX [Streptococcus mutans] Length = 159
SeqID 1693	SA-307.1	Contig126 (20359-20697 m)	45	Identities = 46/150 (30%), Positives = 73/150 (48%), Gaps = 16/150 (10%) dbj BAB07445.1 (AP001519) unknown conserved protein in others [Bacillus halodurans] Length = 152
SeqID 1694	SA-308.1	Contig126 (19077-20267 m)	76	Identities = 243/393 (61%), Positives = 303/393 (76%), Gaps = 8/393 (2%) gb AAD00285.1 (U78604) putative membrane protein [Streptococcus mutans] Length = 395
SeqID 1695	SA-309.1	Contig126 (17861-19102 m)	49	Identities = 133/347 (38%), Positives = 207/347 (59%), Gaps = 5/347 (1%) gb AAA25160.1 (L16975) ORF1 [Lactococcus lactis] Length = 349
SeqID 1696	SA-31.1	Contig137 (11471-11818 p)	No Hits found	
SeqID 1697	SA-311.1	Contig126 (16069-17751 m)	82	Identities = 395/559 (70%), Positives = 465/559 (82%), Gaps = 8/559 (1%) gb AAA25161.1 (L16975) alpha-acetolactate synthase [Lactococcus lactis] Length = 554
SeqID 1698	SA-312.1	Contig126 (15336-16055 m)	72	Identities = 139/239 (58%), Positives = 186/239 (77%), Gaps = 3/239 (1%) gb AAB37482.1 (S82499) alpha-acetolactate decarboxylase, AldB [Lactococcus lactis, ssp. lactis, NCDO2118, Peptide, 236 aa] gb AAB81923.1 (U92974) AldB [Lactococcus lactis] Length = 236
SeqID 1699	SA-313.2	Contig126 (13627-15282 p)	82	Identities = 393/550 (71%), Positives = 462/550 (83%) emb CAA46282.1 (X65164) fibronectin-binding protein-like protein A [Streptococcus gordonii] Length = 550

SeqID 1700	SA-314.2	Contig126 (12265-13233 m)	54	Identities = 112/295 (37%), Positives = 184/295 (61%), Gaps = 1/295 (0%) pir H82240 conserved hypothetical protein VC1101 [imported] - <i>Vibrio cholerae</i> (group O1 strain N16961) gb AAF94260.1 (AE004191) conserved hypothetical protein [Vibrio cholerae] Length = 321
SeqID 1701	SA-315.1	Contig126 (11389-12252 m)	63	Identities = 117/290 (40%), Positives = 186/290 (63%), Gaps = 9/290 (3%) pir F83165 probable permease of ABC transporter PA3837 [imported] - <i>Pseudomonas aeruginosa</i> (strain PAO1) gb AAG07224.1 AE004801_2 (AE004801) probable permease of ABC transporter [Pseudomonas aeruginosa] Length = 296
SeqID 1702	SA-316.1	Contig126 (10628-11389 m)	60	Identities = 116/254 (45%), Positives = 167/254 (65%) gb AAF86640.1 AF162694_1 (AF162694) ABC transporter [Enterococcus gallinarum] Length = 269
SeqID 1703	SA-318.1	Contig126 (8670-10331 m)	64	Identities = 237/555 (42%), Positives = 363/555 (64%), Gaps = 2/555 (0%) dbj BAB06117.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 555
SeqID 1704	SA-319.1	Contig126 (7825-8613 m)	63	Identities = 155/262 (59%), Positives = 189/262 (71%), Gaps = 4/262 (1%) gb AAF62859.1 AF157484_1 (AF157484) tributyrin esterase [Lactococcus lactis subsp. lactis] Length = 258
SeqID 1705	SA-32.1	Contig137 (11161-11478 p)	No Hits found	
SeqID 1706	SA-320.1	Contig126 (7025-7684 m)	No Hits found	
SeqID 1707	SA-321.1	Contig126 (6333-6878 p)	21	Identities = 30/60 (50%), Positives = 39/60 (65%), Gaps = 4/60 (6%) pir F72654 hypothetical protein APE0666 - <i>Aeropyrum pernix</i> (strain K1) dbj BAA79638.1 (AP000060) 102aa long hypothetical protein [Aeropyrum pernix] Length = 102

SeqID 1708	SA-322.1	Contig126 (6116-6787 m)	58	Identities = 103/220 (46%), Positives = 141/220 (63%), Gaps = 3/220 (1%) sp P72012 RPIA_METH_PROBABLE_RIBOSE_5-PHOSPHATE_ISOMERASE (PHOSPHORIBOISOMERASE) pir G69180 ribose 5-phosphate isomerase - Methanobacterium thermoautotrophicum (strain Delta H) dbj BAA13646.1 (D88555) orf2 [Methanobacterium thermoautotrophicum] gb AAB85114.1 (AE000842) ribose 5-phosphate isomerase [Methanobacterium thermoautotrophicum] Length = 226
SeqID 1709	SA-323.1	Contig126 (4848-6059 m)	79	Identities = 275/408 (67%), Positives = 325/408 (79%), Gaps = 7/408 (1%) sp O32808 DEOB_LACLC PHOSPHOPENTOMUTASE (PHOSPHODEOXYRIBOMUTASE) gb AAC45496.1 (U80410) phosphopentomutase [Lactococcus lactis subsp. cremoris] Length = 411
SeqID 1710	SA-324.1	Contig126 (4390-4800 m)	55	Identities = 73/130 (56%), Positives = 93/130 (71%) sp P52147 ARC2_ECOLI_ARSENATE_REDUCTASE (ARSENICAL PUMP MODIFIER) gb AAB09628.1 (U38947) ArsC [Plasmid R46] Length = 141
SeqID 1711	SA-325.1	Contig126 (3542-4351 m)	74	Identities = 149/272 (54%), Positives = 201/272 (73%), Gaps = 3/272 (1%) dbj BAB05250.1 (AP001512) purine nucleoside phosphorylase [Bacillus halodurans] Length = 272
SeqID 1712	SA-326.1	Contig126 (2284-3540 m)	50	Identities = 121/410 (29%), Positives = 215/410 (51%), Gaps = 19/410 (4%) gb AAD53928.1 AF179611_12 (AF179611) chloride channel protein [Zymomonas mobilis] Length = 425
SeqID 1713	SA-327.1	Contig126 (1590-2300 m)	82	Identities = 177/216 (81%), Positives = 197/216 (90%) sp Q56037 DEOD_STRTR_PURINE_NUCLEOSIDE PHOSPHORYLASE (INOSINE PHOSPHORYLASE) (PNP) gb AAC44007.1 (U40830) DeoD [Streptococcus thermophilus] prf 2209356A deoD gene [Streptococcus thermophilus] Length = 216
SeqID 1714	SA-328.1	Contig126 (814-1581 m)	No Hits found	

SeqID 1715	SA-329.1	Contig126 (169-789 p)	99	Identities = 262/262 (100%), Positives = 262/262 (100%) pir T44638 capsular polysaccharide biosynthesis protein cpsY [imported] - Streptococcus agalactiae emb CAB36980.1 (Y17218) CpsY protein [Streptococcus agalactiae] emb CAB36982.2 (Y17241) CpsY protein [Streptococcus agalactiae] gb AAD53064.1 AF163833_1 (AF163833) CpsY [Streptococcus agalactiae] Length = 307
SeqID 1716	SA-330.2	Contig132 (27403-27660 p)	No Hits found	
SeqID 1717	SA-331.2	Contig132 (26320-27393 p)	69	Identities = 212/347 (61%), Positives = 261/347 (75%), Gaps = 3/347 (0%) emb CAA04376.1 (AJ000883) purK [Lactococcus lactis] Length = 349
SeqID 1718	SA-333.1	Contig132 (25842-26333 p)	58	Identities = 102/158 (64%), Positives = 129/158 (81%) emb CAA04375.1 (AJ000883) purE [Lactococcus lactis] Length = 161
SeqID 1719	SA-334.1	Contig132 (24299-25561 p)	68	Identities = 239/419 (57%), Positives = 300/419 (71%), Gaps = 7/419 (1%) sp Q9ZF44 PUR2_LACLA PHOSPHORIBOSYLAMINE--GLYCINE LIGASE (GARS) (GLYCINAMIDE RIBONUCLEOTIDE SYNTHETASE) (PHOSPHORIBOSYLGLYCINAMIDE SYNTHETASE) emb CAA04374.1 (AJ000883) purD [Lactococcus lactis] Length = 412
SeqID 1720	SA-335.1	Contig132 (23212-24015 m)	41	Identities = 66/258 (25%), Positives = 119/258 (45%), Gaps = 9/258 (3%) sp P26833 YNGB_CLOPE HYPOTHETICAL 31.2 KDA PROTEIN IN NAGH 5 REGION (ORFB) pir S43902 hypothetical protein B - Clostridium perfringens gb AA23257.1 (M81878) unknown [Clostridium perfringens] Length = 279
SeqID 1721	SA-336.1	Contig132 (22212-23189 p)	55	Identities = 133/299 (44%), Positives = 188/299 (62%), Gaps = 1/299 (0%) pir S41858 hypothetical protein - Thermoanaerobacterium saccharolyticum Length = 320

SeqID 1722	SA-337.1	Contig132 (21323-22204 p)	37	Identities = 78/160 (48%), Positives = 110/160 (68%) sp P26832 YNGA_CLOPE HYPOTHETICAL PROTEIN IN NAGH 5 REGION (ORFA) pir S43901 hypothetical protein A - Clostridium perfringens (fragment) gb AA23256.1 (M81878) unknown [Clostridium perfringens] Length = 182
SeqID 1723	SA-338.1	Contig132 (20389-21306 p)	58	Identities = 162/225 (72%), Positives = 191/225 (84%) emb CAA69950.1 (Y08695) putative acylneuraminase [Clostridium tertium] Length = 226
SeqID 1724	SA-339.1	Contig132 (19703-20392 p)	32	Identities = 40/148 (27%), Positives = 74/148 (49%), Gaps = 4/148 (2%) dbj BAB05827.1 (AP001514) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 214
SeqID 1725	SA-34.1	Contig137 (10492-10767 p)	40	Identities = 30/91 (32%), Positives = 50/91 (53%), Gaps = 6/91 (6%) gb AAF73773.1 AF154006_1 (AF154006) surface protein PspC [Streptococcus pneumoniae] Length = 678
SeqID 1726	SA-340.1	Contig132 (19267-19710 p)	43	Identities = 48/151 (31%), Positives = 66/151 (42%), Gaps = 5/151 (3%) gb AAC44392.1 (U43526) ORF-1 [Streptococcus pneumoniae] Length = 150
SeqID 1727	SA-342.1	Contig132 (18424-19254 p)	56	Identities = 94/262 (35%), Positives = 158/262 (59%), Gaps = 1/262 (0%) pir F72379 sugar ABC transporter, permease protein Thermotoga maritima (strain MSB8) gb AAD35515.1 AE001721_14 (AE001721) sugar ABC transporter, permease protein [Thermotoga maritima] Length = 271
SeqID 1728	SA-343.1	Contig132 (17527-18414 p)	51	Identities = 106/289 (36%), Positives = 168/289 (57%), Gaps = 6/289 (2%) dbj BAB05584.1 (AP001513) sugar transport system (permease) (binding protein dependent transporter) [Bacillus halodurans] Length = 309
SeqID 1729	SA-344.1	Contig132 (16123-17439 p)	20	Identities = 54/187 (28%), Positives = 90/187 (47%), Gaps = 14/187 (7%) pir F69796 sugar-binding protein homolog yesO - Bacillus subtilis emb CAB12516.1 (Z99107) similar to sugar- binding protein [Bacillus subtilis] Length = 412

SeqID 1730	SA-345.1	Contig132 (15378-16076 p)	62	Identities = 107/226 (47%), Positives = 148/226 (65%), Gaps = 6/226 (2%) pir C70180 conserved hypothetical protein BB0644 - Lyme disease spirochete gb AAC66999.1 (AE001166) conserved hypothetical protein [Borrelia burgdorferi] Length = 232
SeqID 1731	SA-347.1	Contig132 (13827-15131 p)	97	Identities = 434/434 (100%), Positives = 434/434 (100%) gb AAG18476.1 (AF151359) group B streptococcal surface immunogenic protein [Streptococcus agalactiae] gb AAG18477.1 (AF151361) group B streptococcal surface immunogenic protein [Streptococcus agalactiae] Length = 434
SeqID 1732	SA-348.1	Contig132 (12781-13680 p)	62	Identities = 163/274 (59%), Positives = 196/274 (71%), Gaps = 11/274 (4%) gb AAC46072.1 (U50357) zoocin A endopeptidase [Streptococcus zooepidemicus] Length = 285
SeqID 1733	SA-351.1	Contig132 (11041-12588 p)	74	Identities = 310/515 (60%), Positives = 390/515 (75%), Gaps = 4/515 (0%) dbj BAB04352.1 (AP001509) phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase [Bacillus halodurans] Length = 511
SeqID 1734	SA-352.1	Contig132 (10269-11021 p)	44	Identities = 65/263 (24%), Positives = 117/263 (43%), Gaps = 34/263 (12%) pir H69779 antibiotic resistance protein homolog ydfB - Bacillus subtilis dbj BAA19369.1 (AB001488) SIMILAR TO BACILLUS CEREUS ZWITTERMICIN A-RESISTANCE GENE. [Bacillus subtilis] emb CAB12342.1 (Z99106) similar to antibiotic resistance protein [Bacillus subtilis] Length = 261
SeqID 1735	SA-353.1	Contig132 (9698-10246 p)	61	Identities = 137/152 (90%), Positives = 140/152 (91%), Gaps = 2/152 (1%) gb AAF08602.1 U70775_1 (U70775) phosphoribosylglycinamide formyltransferase homolog [Streptococcus pyogenes] Length = 151

SeqID 1736	SA-356.1	Contig132 (8508-9530 p)	85	Identities = 257/339 (75%), Positives = 293/339 (85%), Gaps = 4/339 (1%) gb AAC16901.1 (AF016634) phosphoribosylformylglycinamide cyclo-ligase [Lactococcus lactis subsp. cremoris] Length = 340
SeqID 1737	SA-357.1	Contig132 (7026-8480 p)	83	Identities = 343/470 (72%), Positives = 406/470 (85%), Gaps = 6/470 (1%) pir T51702 amidophosphoribosyltransferase (EC 2.4.2.14) [imported] - Lactococcus lactis gb AAD12627.1 (U64311) phosphoribosylpyrophosphate amidotransferase [Lactococcus lactis] Length = 506
SeqID 1738	SA-36.1	Contig137 (9548-10369 p)	No Hits found	
SeqID 1739	SA-361.1	Contig132 (4533-5015 m)	39	Identities = 39/137 (28%), Positives = 63/137 (45%), Gaps = 7/137 (5%) prf 2119294A YFW1 gene [Saccharomyces cerevisiae] Length = 605
SeqID 1740	SA-363.1	Contig132 (3046-6792 p)	25	Identities = 198/746 (26%), Positives = 330/746 (43%), Gaps = 103/746 (13%) emb CAC12194.1 (AL445066) phosphoribosylformylglycinamide synthase related protein [Thermoplasma acidophilum] Length = 759
SeqID 1741	SA-364.2	Contig132 (2240-2944 p)	84	Identities = 183/235 (77%), Positives = 206/235 (86%) sp Q07296 PUR7_STRPN PHOSPHORIBOSYLAMINOIMIDAZOLE- SUCCINOCARBOXAMIDE SYNTHASE (SAICAR SYNTHETASE) pir A36941 phosphoribosylaminoimidazolesuccinocarboxamide synthase (EC 6.3.2.6) - Streptococcus pneumoniae gb AAA03540.1 (L15190) SAICAR synthetase [Streptococcus pneumoniae] gb AAA69512.1 (M36180) SAICAR synthetase [Streptococcus pneumoniae] Length = 235
SeqID 1742	SA-366.2	Contig139 (35877-38138 m)	85	Identities = 560/755 (74%), Positives = 648/755 (85%), Gaps = 12/755 (1%) gb AAD01782.1 (AF023421) ClpE [Lactococcus lactis] Length = 748

SeqID 1743	SA-367.1	Contig139 (35236-35691 m)	29	Identities = 31/101 (30%), Positives = 48/101 (46%), Gaps = 2/101 (1%) pir A70315 AP4A hydrolase - Aquifex aeolicus gb AAC06510.1 (AE000676) AP4A hydrolase [Aquifex aeolicus] Length = 134
SeqID 1744	SA-368.1	Contig139 (34870-35172 m)	No Hits found	
SeqID 1745	SA-37.1	Contig137 (8461-9465 p)	37	Identities = 63/288 (21%), Positives = 129/288 (43%), Gaps = 27/288 (9%) ref NP_053160.1 pXO2-05 [Bacillus anthracis] gb AAF13610.1 AF188935_7 (AF188935) pXO2-05 [Bacillus anthracis] Length = 282
SeqID 1746	SA-370.1	Contig139 (31961-34753 p)	87	Identities = 729/929 (78%), Positives = 821/929 (87%), Gaps = 1/929 (0%) sp Q9ZHB3 SYL_STRPN ISOLEUCYL-TRNA SYNTHETASE (ISOLEUCINE--TRNA LIGASE) (ILERS) gb AAC95446.1 (AF068901) isoleucine-tRNA synthetase [Streptococcus pneumoniae] Length = 930
SeqID 1747	SA-371.1	Contig139 (30906-31676 p)	68	Identities = 132/227 (58%), Positives = 179/227 (78%), Gaps = 2/227 (0%) gb AAC95445.1 (AF068901) cell division protein DivIVA [Streptococcus pneumoniae] Length = 262
SeqID 1748	SA-372.1	Contig139 (30108-30896 p)	60	Identities = 101/255 (39%), Positives = 162/255 (62%) gb AAC95444.1 (AF068901) YlmH [Streptococcus pneumoniae] Length = 261
SeqID 1749	SA-373.1	Contig139 (29852-30106 p)	59	Identities = 34/83 (40%), Positives = 54/83 (64%) emb CAA75619.1 (Y15422) hypothetical protein [Lactococcus lactis subsp. cremoris] Length = 91
SeqID 1750	SA-374.1	Contig139 (29244-29849 p)	57	Identities = 86/200 (43%), Positives = 120/200 (60%), Gaps = 25/200 (12%) gb AAC95442.1 (AF068901) YlmF [Streptococcus pneumoniae] Length = 179
SeqID 1751	SA-375.1	Contig139 (28558-29232 p)	76	Identities = 140/223 (62%), Positives = 177/223 (78%) gb AAC95441.1 (AF068901) YlmE [Streptococcus pneumoniae] Length = 223
SeqID 1752	SA-376.1	Contig139 (27272-28552 p)	84	Identities = 327/426 (76%), Positives = 363/426 (84%), Gaps = 7/426 (1%) gb AAC95440.1 (AF068901) cell division protein FtsZ [Streptococcus pneumoniae] Length = 419

SeqID 1753	SA-377.1	Contig139 (25877-27250 p)	80	Identities = 293/458 (63%), Positives = 368/458 (79%), Gaps = 1/458 (0%) gb AAC95439.1 (AF068901) cell division protein FtsA [Streptococcus pneumoniae] Length = 457
SeqID 1754	SA-379.1	Contig139 (24468-25604 p)	55	Identities = 123/396 (31%), Positives = 210/396 (52%), Gaps = 38/396 (9%) gb AAC95451.1 (AF068902) cell division protein DivIB [Streptococcus pneumoniae] Length = 399
SeqID 1755	SA-380.1	Contig139 (23388-24464 p)	80	Identities = 248/358 (69%), Positives = 293/358 (81%), Gaps = 1/358 (0%) dbj BAB16029.1 (AB030645) MurG [Streptococcus pyogenes] Length = 360
SeqID 1756	SA-381.1	Contig139 (22030-23385 p)	84	Identities = 340/449 (75%), Positives = 392/449 (86%) gb AAC95449.1 (AF068902) D-glutamic acid adding enzyme MurD [Streptococcus pneumoniae] Length = 450
SeqID 1757	SA-382.1	Contig139 (21655-21900 p)	No Hits found	
SeqID 1758	SA-384.1	Contig139 (19769-21610 p)	84	Identities = 457/609 (75%), Positives = 536/609 (87%), Gaps = 2/609 (0%) sp O07631 TYPA_BACSU GTP-BINDING PROTEIN TYPA/BIPA HOMOLOG pir E69872 GTP-binding translation elongation factor homolog ylaG - Bacillus subtilis emb CAB09712.1 (Z97025) product highly similar to elongation factor EF-G [Bacillus subtilis] emb CAB13350.1 (Z99111) similar to GTP-binding elongation factor [Bacillus subtilis] Length = 612
SeqID 1759	SA-385.1	Contig139 (19157-19537 p)	55	Identities = 51/124 (41%), Positives = 71/124 (57%), Gaps = 1/124 (0%) sp P54510 YQHL_BACSU HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIA INTERGENIC REGION pir C69959 glpE protein homolog yqhl - Bacillus subtilis dbj BAA12549.1 (D84432) Yqhl [Bacillus subtilis] emb CAB14385.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] Length = 126

SeqID 1760	SA-387.1	Contig139 (18177-19145 p)	64	Identities = 145/315 (46%), Positives = 209/315 (66%), Gaps = 2/315 (0%) dbj BAB05144.1 (AP001512) glucose kinase [Bacillus halodurans] Length = 330
SeqID 1761	SA-388.1	Contig139 (17971-18180 p)	84	Identities = 53/67 (79%), Positives = 62/67 (92%) dbj BAA96473.1 (AB036428) hypothetical 8.3 kDa protein [Streptococcus mutans] Length = 68
SeqID 1762	SA-389.1	Contig139 (17421-17867 m)	57	Identities = 57/123 (46%), Positives = 85/123 (68%) dbj BAA96471.1 (AB036428) type IV prepilin peptidase homologue [Streptococcus mutans] Length = 218
SeqID 1763	SA-39.1	Contig137 (7426-8382 m)	No Hits found	
SeqID 1764	SA-390.1	Contig139 (16788-17420 p)	61	Identities = 95/202 (47%), Positives = 134/202 (66%) dbj BAB05417.1 (AP001512) endonuclease III (DNA repair) [Bacillus halodurans] Length = 218
SeqID 1765	SA-391.1	Contig139 (15458-16675 p)	38	Identities = 88/384 (22%), Positives = 159/384 (40%), Gaps = 24/384 (6%) pir A69832 probable acid-CoA ligase (EC 6.2.1.-) yhfT [similarity] - Bacillus subtilis emb CAA74543.1 (Y14084) hypothetical protein [Bacillus subtilis] emb CAB12876.1 (Z99109) similar to long-chain fatty-acid-CoA ligase [Bacillus subtilis] Length = 479
SeqID 1766	SA-392.1	Contig139 (14410-15453 p)	45	Identities = 96/340 (28%), Positives = 160/340 (46%), Gaps = 21/340 (6%) emb CAC11722.1 (AL445064) acetyl-CoA acetyltransferase related protein [Thermoplasma acidophilum] Length = 388
SeqID 1767	SA-393.1	Contig139 (13834-14349 p)	No Hits found	
SeqID 1768	SA-394.1	Contig139 (12717-13709 m)	55	Identities = 132/316 (41%), Positives = 201/316 (62%), Gaps = 2/316 (0%) dbj BAB05467.1 (AP001513) biotin synthase [Bacillus halodurans] Length = 333

SeqID 1769	SA-395.1	Contig139 (12177-12716 m)	58	Identities = 69/168 (41%), Positives = 105/168 (62%) pir B69832 biotin biosynthesis homolog yhfU - Bacillus subtilis emb CAA74544.1 (Y14084) hypothetical protein [Bacillus subtilis] emb CAB12877.1 (Z99109) similar to biotin biosynthesis [Bacillus subtilis] Length = 186
SeqID 1770	SA-396.1	Contig139 (11543-12109 p)	60	Identities = 102/191 (53%), Positives = 134/191 (69%), Gaps = 4/191 (2%) sp Q02003 TRPG_LACLA ANTHRANILATE SYNTHASE COMPONENT II (GLUTAMINE AMIDO-TRANSFERASE) pir S35125 anthranilate synthase (EC 4.1.3.27) beta chain - Lactococcus lactis subsp. lactis gb AAA25224.1 (M87483) anthranilate synthase beta subunit [Lactococcus lactis] Length = 198
SeqID 1771	SA-397.1	Contig139 (9673-11415 p)	68	Identities = 280/582 (48%), Positives = 400/582 (68%), Gaps = 6/582 (1%) dbj BAB06054.1 (AP001515) ABC transporter (ATP-binding protein) [Bacillus halodurans] Length = 599
SeqID 1772	SA-398.1	Contig139 (7941-9686 p)	69	Identities = 284/575 (49%), Positives = 406/575 (70%), Gaps = 2/575 (0%) dbj BAB06055.1 (AP001515) ABC transporter (ATP-binding protein) [Bacillus halodurans] Length = 585
SeqID 1773	SA-399.1	Contig139 (7161-7745 p)	34	Identities = 46/154 (29%), Positives = 68/154 (43%), Gaps = 36/154 (23%) gb AAF36806.1 (AF155139) VanZF [Paenibacillus popilliae] Length = 206
SeqID 1774	SA-4.1	Contig137 (39841-40488 p)	No Hits found	
SeqID 1775	SA-40.1	Contig137 (7185-7433 p)	No Hits found	
SeqID 1776	SA-400.1	Contig139 (6106-7230 p)	67	Identities = 206/349 (59%), Positives = 257/349 (73%), Gaps = 5/349 (1%) dbj BAB06225.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 362
SeqID 1777	SA-401.2	Contig139 (5495-6073 p)	29	Identities = 46/89 (51%), Positives = 62/89 (68%), Gaps = 1/89 (1%) pir F70023 hypothetical protein yutD - Bacillus subtilis emb CAB15221.1 (Z99120) yutD [Bacillus subtilis] Length = 102

SeqID 1778	SA-402.2	Contig139 (4092-5426 p)	58	Identities = 155/463 (33%), Positives = 262/463 (56%), Gaps = 33/463 (7%) pir G70015 conserved hypothetical protein yunD - Bacillus subtilis emb CAB15227.1 (Z99120) similar to hypothetical proteins [Bacillus subtilis] Length = 462
SeqID 1779	SA-403.1	Contig139 (2944-4017 p)	55	Identities = 125/344 (36%), Positives = 199/344 (57%), Gaps = 21/344 (6%) pir A69875 hypothetical protein ybL - Bacillus subtilis emb CAB11358.1 (Z98682) YbL protein [Bacillus subtilis] emb CAB13378.1 (Z99111) ybL [Bacillus subtilis] Length = 350
SeqID 1780	SA-405.1	Contig139 (2469-2954 p)	57	Identities = 56/149 (37%), Positives = 94/149 (62%) sp P57643 COAD_BUCAI PHOSPHOPANTETHEINE ADENYLTRANSFERASE (PANTETHEINE-PHOSPHATE ADENYLTRANSFERASE) (PPAT) (DEPHOSPHO-COA PYROPHOSPHORYLASE) dbj BAB13272.1 (AP001119) lipopolysaccharide core biosynthesis protein kdtB [Buchnera sp. APS] Length = 165
SeqID 1781	SA-407.1	Contig139 (2149-2472 p)	No Hits found	
SeqID 1782	SA-409.1	Contig139 (1418-2170 p)	46	Identities = 80/180 (44%), Positives = 118/180 (65%), Gaps = 3/180 (1%) dbj BAB06309.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 189
SeqID 1783	SA-41.1	Contig137 (7096-7170 p)	No Hits found	
SeqID 1784	SA-410.1	Contig139 (1051-1419 m)	31	Identities = 30/97 (30%), Positives = 51/97 (51%), Gaps = 3/97 (3%) gb AAD56628.1 AF165218_3 (AF165218) Bta [Streptococcus pneumoniae] Length = 115
SeqID 1785	SA-411.1	Contig139 (5-1006 p)	86	Identities = 267/328 (81%), Positives = 290/328 (88%) sp P44338 ASNA_HAEIN ASPARTATE--AMMONIA LIGASE (ASPARAGINE SYNTHETASE A) pir H64077 aspartate-- ammonia ligase (EC 6.3.1.1) - Haemophilus influenzae (strain Rd KW20) gb AAC22222.1 (U32738) aspartate-- ammonia ligase (asnA) [Haemophilus influenzae Rd] Length = 330

SeqID 1786	SA-412.2	Contig134 (36525-37172 m)	75	Identities = 125/212 (58%), Positives = 165/212 (76%) gb AAF81675.1 AF232688_4 (AF232688) SloR [Streptococcus mutans] Length = 217
SeqID 1787	SA-413.1	Contig134 (35796-36485 p)	62	Identities = 100/229 (43%), Positives = 145/229 (62%) sp P45113 MTN_HAEIN_MTA/SAH_NUCLEOSIDASE [INCLUDES: 5-METHYLTHIOADENOSINE NUCLEOSIDASE ; S-ADENOSYLHOMOCYSTEINE NUCLEOSIDASE] pir C64169 pfs protein - Haemophilus influenzae (strain Rd KW20) gb AAC22869.1 (U32801) pfs protein (pfs) [Haemophilus influenzae Rd] Length = 229
SeqID 1788	SA-414.1	Contig134 (35517-35786 p)	No Hits found	
SeqID 1789	SA-415.1	Contig134 (34963-35517 p)	62	Identities = 92/181 (50%), Positives = 125/181 (68%), Gaps = 4/181 (2%) sp P54570 YQKG_BACSU_HYPOTHETICAL 21.0 KDA PROTEIN IN GLNQ-ANSR INTERGENIC REGION pir A69967 conserved hypothetical protein yqgG - Bacillus subtilis dbj BAA12639.1 (D84432) YqgG [Bacillus subtilis] emb CAB14293.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] Length = 185
SeqID 1790	SA-416.1	Contig134 (33563-34942 p)	68	Identities = 238/453 (52%), Positives = 322/453 (70%), Gaps = 1/453 (0%) dbj BAB03784.1 (AP001507) UDP-N- acetylglucosamine pyrophosphorylase [Bacillus halodurans] Length = 455
SeqID 1791	SA-417.1	Contig134 (32906-33319 p)	58	Identities = 46/144 (31%), Positives = 83/144 (56%), Gaps = 10/144 (6%) dbj BAB04569.1 (AP001510) unknown conserved protein in others [Bacillus halodurans] Length = 148
SeqID 1792	SA-418.1	Contig134 (32516-32893 p)	48	Identities = 41/127 (32%), Positives = 63/127 (49%), Gaps = 11/127 (8%) pir E69798 conserved hypothetical protein yeth - Bacillus subtilis emb CAB12535.1 (Z99107) similar to hypothetical proteins [Bacillus subtilis] Length = 120
SeqID 1793	SA-419.1	Contig134 (31471-32427 p)	60	Identities = 133/320 (41%), Positives = 193/320 (59%), Gaps = 4/320 (1%) dbj BAB06422.1 (AP001516) unknown conserved protein [Bacillus halodurans] Length = 329

SeqID 1794	SA-42.1	Contig137 (6439-6849 p)	No Hits found	
SeqID 1795	SA-420.1	Contig134 (31220-31474 p)	50	Identities = 27/90 (30%), Positives = 49/90 (54%), Gaps = 7/90 (7%) pir T35570 hypothetical protein SC6G4.19c SC6G4.19c - Streptomyces coelicolor emb CAA20397.1 (AL031317) SC6G4.19c, unknown, len: 190 aa; contains Pro-Ser-rich domain at N-terminus [Streptomyces coelicolor A3(2)] Length = 190
SeqID 1796	SA-421.1	Contig134 (31045-31359 m)	No Hits found	
SeqID 1797	SA-422.1	Contig134 (31017-31259 p)	No Hits found	
SeqID 1798	SA-423.1	Contig134 (30787-31020 p)	No Hits found	
SeqID 1799	SA-424.1	Contig134 (30080-30778 p)	52	Identities = 82/239 (34%), Positives = 125/239 (51%), Gaps = 15/239 (6%) pir T44434 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) [imported] - Moritella marina dbj BAA85256.1 (AB021978) 3-oxoacyl-[acyl carrier protein] reductase homolog [Moritella marina] Length = 244
SeqID 1800	SA-425.1	Contig134 (29851-30087 p)	No Hits found	
SeqID 1801	SA-426.1	Contig134 (29501-29761 p)	No Hits found	
SeqID 1802	SA-427.1	Contig134 (28815-29315 p)	33	Identities = 35/153 (22%), Positives = 64/153 (40%), Gaps = 5/153 (3%) pir T35964 hypothetical protein SC9C7.13c - Streptomyces coelicolor emb CAA22725.1 (AL035161) hypothetical protein SC9C7.13c [Streptomyces coelicolor A3(2)] Length = 179
SeqID 1803	SA-428.1	Contig134 (28120-28575 p)	44	Identities = 30/138 (21%), Positives = 68/138 (48%) dbj BAB05950.1 (AP001514) unknown conserved protein in others [Bacillus halodurans] Length = 482
SeqID 1804	SA-429.1	Contig134 (27217-28119 p)	55	Identities = 126/320 (39%), Positives = 180/320 (55%), Gaps = 18/320 (5%) dbj BAB05950.1 (AP001514) unknown conserved protein in others [Bacillus halodurans] Length = 482

SeqID 1805	SA-430.1	Contig134 (24924-27023 p)	44	Identities = 199/697 (28%), Positives = 322/697 (45%), Gaps = 58/697 (8%) dbj BAB05949.1 (AP001514) unknown [Bacillus halodurans] Length = 1091
SeqID 1806	SA-431.1	Contig134 (23373-24857 p)	29	Identities = 71/359 (19%), Positives = 148/359 (40%), Gaps = 49/359 (13%) sp Q45535 COTH_BACSU INNER SPORE COAT PROTEIN H pir E69605 spore coat protein (inner) coth - Bacillus subtilis emb CAB07793.1 (Z93767) ywrH [Bacillus subtilis] emb CAB15623.1 (Z99122) spore coat protein (inner) [Bacillus subtilis] Length = 362
SeqID 1807	SA-432.1	Contig134 (22742-23419 p)	No Hits found	
SeqID 1808	SA-433.1	Contig134 (22079-22690 p)	No Hits found	
SeqID 1809	SA-434.2	Contig134 (20288-21631 p)	55	Identities = 162/436 (37%), Positives = 248/436 (56%), Gaps = 13/436 (2%) pir C69596 branched-chain amino acid transporter braB - Bacillus subtilis gb AAC00400.1 (AF008220) branch-chain amino acid transporter [Bacillus subtilis] emb CAB14938.1 (Z99119) branched-chain amino acid transporter [Bacillus subtilis] Length = 445
SeqID 1810	SA-435.1	Contig134 (18101-20098 p)	72	Identities = 395/667 (59%), Positives = 501/667 (74%), Gaps = 12/667 (1%) sp P37465 SYM_BACSU METHIONYL-TRNA SYNTHETASE (METHIONINE--TRNA LIGASE) (METRS) pir S66067 methionine--trna ligase (EC 6.1.1.10) metS - Bacillus subtilis dbj BAA05273.1 (D26185) methionyl-tRNA synthetase [Bacillus subtilis] emb CAB11814.1 (Z99104) methionyl-tRNA synthetase [Bacillus subtilis] Length = 664
SeqID 1811	SA-436.1	Contig134 (17083-17958 m)	69	Identities = 163/282 (57%), Positives = 205/282 (71%), Gaps = 1/282 (0%) sp P45134 TEHB_HAEIN TELLURITE RESISTANCE PROTEIN TEHB HOMOLOG pir H64113 hemagglutinin hag1 homolog - Haemophilus influenzae (strain Rd KW20) gb AAC22923.1 (U32807) tellurite resistance protein (tehB) [Haemophilus influenzae Rd] Length = 286

SeqID 1812	SA-437.1	Contig134 (16203-16925 p)	No Hits found	
SeqID 1813	SA-438.1	Contig134 (14677-15894 m)	19	Identities = 60/197 (30%), Positives = 83/197 (41%), Gaps = 12/197 (6%) sp P17334 PTCC_ECOLI_PTS_SYSTEM, CELLOBIOSE-SPECIFIC IIC COMPONENT (EIIIC-CEL) (CELLOBIOSE-PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT) pir A64933 celB protein - Escherichia coli (strain K-12) gb AAC74807.1 (AE000268) PEP-dependent phosphotransferase enzyme II for cellobiose, arbutin, and salicin [Escherichia coli K12] Length = 452
SeqID 1814	SA-439.1	Contig134 (13851-14636 m)	33	Identities = 53/240 (22%), Positives = 102/240 (42%), Gaps = 24/240 (10%) gb AAD11512.1 (U60828) unknown [Lactococcus lactis] Length = 307
SeqID 1815	SA-44.1	Contig137 (5997-6446 p)	No Hits found	
SeqID 1816	SA-441.1	Contig134 (12947-13774 m)	87	Identities = 218/275 (79%), Positives = 246/275 (89%) sp P21998 EXOA_STRPN_EXODEOXYRIBONUCLEASE pir A32301 exodeoxyribonuclease (EC 3.1.11.-) exoA - Streptococcus pneumoniae gb AA26879.1 (J04234) exodeoxyribonuclease [Streptococcus pneumoniae] Length = 275
SeqID 1817	SA-442.1	Contig134 (12550-12906 p)	58	Identities = 56/107 (52%), Positives = 74/107 (68%), Gaps = 1/107 (0%) dbj BAB07204.1 (AP001518) arsenate reductase [Bacillus halodurans] Length = 119
SeqID 1818	SA-443.1	Contig134 (12072-12548 p)	57	Identities = 75/156 (48%), Positives = 99/156 (63%), Gaps = 9/156 (5%) pir F82390 methylated-DNA--protein-cysteine S- methyltransferase (EC 2.1.1.63) VCA1017 [similarity] - Vibrio cholerae (group O1 strain N16961) gb AAF96913.1 (AE004427) methylated-DNA--protein-cysteine S- methyltransferase [Vibrio cholerae] Length = 157

SeqID 1819	SA-445.1	Contig134 (10835-12016 p)	42	Identities = 102/313 (32%), Positives = 168/313 (53%), Gaps = 21/313 (6%) sp Q58424 SERA_METJA D-3-PHOSPHOGLYCERATE DEHYDROGENASE (PGDH) pir A64427 phosphoglycerate dehydrogenase (EC 1.1.1.95) - Methanococcus jannaschii gb AAB99020.1 (U67544) phosphoglycerate dehydrogenase (serA) [Methanococcus jannaschii] Length = 524
SeqID 1820	SA-446.1	Contig134 (10213-10773 p)	41	Identities = 45/170 (26%), Positives = 78/170 (45%), Gaps = 13/170 (7%) pir G81269 probable acetyltransferase Cj1715 [imported] - Campylobacter jejuni (strain NCTC 11168) emb CAB73701.1 (AL139079) putative acetyltransferase [Campylobacter jejuni] Length = 176
SeqID 1821	SA-447.1	Contig134 (9065-10156 p)	67	Identities = 169/363 (46%), Positives = 252/363 (68%), Gaps = 8/363 (2%) gb AAAF13453.1 AF204962_1 (AF204962) phosphoserine aminotransferase [Bacillus alcalophilus] Length = 361
SeqID 1822	SA-448.1	Contig134 (8297-8932 m)	54	Identities = 80/203 (39%), Positives = 116/203 (56%), Gaps = 7/203 (3%) ref XP_005810.1 CGI-32 protein [Homo sapiens] Length = 307
SeqID 1823	SA-45.1	Contig137 (5361-5807 p)	No Hits found	
SeqID 1824	SA-450.1	Contig134 (7155-8027 p)	65	Identities = 141/287 (49%), Positives = 190/287 (66%), Gaps = 2/287 (0%) dbj BAB03768.1 (AP001507) unknown conserved protein [Bacillus halodurans] Length = 289
SeqID 1825	SA-451.1	Contig134 (6832-7158 p)	56	Identities = 45/116 (38%), Positives = 62/116 (52%), Gaps = 8/116 (6%) dbj BAB03765.1 (AP001507) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 116
SeqID 1826	SA-452.1	Contig134 (5938-6801 p)	41	Identities = 82/219 (37%), Positives = 124/219 (56%), Gaps = 3/219 (1%) dbj BAB03763.1 (AP001507) DNA polymerase III delta subunit [Bacillus halodurans] Length = 328

SeqID 1827	SA-453.1	Contig134 (5283-5918 p)	59	Identities = 112/210 (53%), Positives = 148/210 (70%), Gaps = 1/210 (0%) dbj BAB03761.1 (AP001507) thymidylate kinase [Bacillus halodurans] Length = 210
SeqID 1828	SA-454.1	Contig134 (4535-5194 p)	51	Identities = 73/166 (43%), Positives = 116/166 (68%), Gaps = 2/166 (1%) pir H72290 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD36216.1 AE001771_9 (AE001771) conserved hypothetical protein [Thermotoga maritima] Length = 215
SeqID 1829	SA-455.1	Contig134 (3806-4516 p)	75	Identities = 135/233 (57%), Positives = 180/233 (76%) pir T35757 probable branched chain amino acid transport ATP-binding protein - Streptomyces coelicolor emb CAB52068.1 (AL109732) putative branched chain amino acid transport ATP-binding protein [Streptomyces coelicolor A3(2)] Length = 238
SeqID 1830	SA-456.1	Contig134 (3042-3806 p)	74	Identities = 136/273 (49%), Positives = 190/273 (68%), Gaps = 21/273 (7%) pir F72290 branched chain amino acid ABC transporter, ATP-binding protein - Thermotoga maritima (strain MSB8) gb AAD36214.1 AE001771_7 (AE001771) branched chain amino acid ABC transporter, ATP-binding protein [Thermotoga maritima] Length = 284
SeqID 1831	SA-457.1	Contig134 (2088-3041 p)	60	Identities = 121/343 (35%), Positives = 195/343 (56%), Gaps = 36/343 (10%) pir E72290 branched chain amino acid ABC transporter, permease protein - Thermotoga maritima (strain MSB8) gb AAD36213.1 AE001771_6 (AE001771) branched chain amino acid ABC transporter, permease protein [Thermotoga maritima] Length = 359

SeqID 1832	SA-458.2	Contig134 (1216-2085 p)	64	Identities = 140/295 (47%), Positives = 200/295 (67%), Gaps = 7/295 (2%) pir D72290 branched chain amino acid ABC transporter, permease protein - <i>Thermotoga maritima</i> (strain MSB8) gb AAD36212.1 AE001771_5 (AE001771) branched chain amino acid ABC transporter, permease protein [<i>Thermotoga maritima</i>] Length = 299
SeqID 1833	SA-46.1	Contig137 (4933-5361 p)	No Hits found	
SeqID 1834	SA-460.2	Contig134 (1-1110 p)	49	Identities = 114/360 (31%), Positives = 182/360 (49%), Gaps = 27/360 (7%) pir C72290 branched chain amino acid ABC transporter, periplasmic amino acid-binding protein - <i>Thermotoga maritima</i> (strain MSB8) gb AAD36211.1 AE001771_4 (AE001771) branched chain amino acid ABC transporter, periplasmic amino acid-binding protein [<i>Thermotoga maritima</i>] Length = 370
SeqID 1835	SA-462.1	Contig135 (65101-65484 m)	40	Identities = 30/102 (29%), Positives = 60/102 (58%) gb AAD05186.1 (AF110185) unknown [<i>Burkholderia pseudomallei</i>] Length = 163
SeqID 1836	SA-463.1	Contig135 (64130-65101 m)	83	Identities = 224/324 (69%), Positives = 276/324 (85%), Gaps = 1/324 (0%) gb AAF98271.1 AF197933_1 (AF197933) beta-ketoacyl-ACP synthase III [<i>Streptococcus pneumoniae</i>] Length = 324
SeqID 1837	SA-464.1	Contig135 (63848-64072 m)	71	Identities = 64/74 (86%), Positives = 67/74 (90%) gb AAF98272.1 AF197933_2 (AF197933) acyl carrier protein [<i>Streptococcus pneumoniae</i>] Length = 74
SeqID 1838	SA-465.1	Contig135 (62734-63693 m)	85	Identities = 257/318 (80%), Positives = 277/318 (86%), Gaps = 1/318 (0%) gb AAF98273.1 AF197933_3 (AF197933) trans-2-enoyl-ACP reductase II [<i>Streptococcus pneumoniae</i>] Length = 324

SeqID 1839	SA-466.1	Contig135 (61788-62714 m)	77	Identities = 204/306 (66%), Positives = 242/306 (78%), Gaps = 1/306 (0%) gb AAAF98274.1 AF197933_4 (AF197933) malonyl-CoA:ACP transacylase; MCAT [Streptococcus pneumoniae] Length = 306
SeqID 1840	SA-467.1	Contig135 (61045-61779 m)	84	Identities = 184/243 (75%), Positives = 212/243 (86%) gb AAAF98275.1 AF197933_5 (AF197933) beta-ketoacyl-ACP reductase [Streptococcus pneumoniae] Length = 243
SeqID 1841	SA-468.1	Contig135 (59797-61029 m)	90	Identities = 340/410 (82%), Positives = 375/410 (90%) gb AAAF98276.1 AF197933_6 (AF197933) beta-ketoacyl-ACP synthase II [Streptococcus pneumoniae] Length = 411
SeqID 1842	SA-469.1	Contig135 (59295-59795 m)	58	Identities = 103/169 (60%), Positives = 127/169 (74%), Gaps = 11/169 (6%) gb AAAF98277.1 AF197933_7 (AF197933) biotin carboxyl carrier protein [Streptococcus pneumoniae] Length = 161
SeqID 1843	SA-47.1	Contig137 (2969-4462 p)	73	Identities = 278/497 (55%), Positives = 367/497 (72%), Gaps = 1/497 (0%) pir JH0206 hypothetical 57.4K protein - Enterococcus faecalis plasmid pAM-beta-1 gb AAB94754.1 (AF039139) replication protein E [Cloning vector pL252] gb AAB96788.1 (AF041239) replication protein E [Cloning vector pL253] gb AAC38603.1 (AF007787) RepE [Enterococcus faecalis] Length = 496
SeqID 1844	SA-470.1	Contig135 (58876-59298 m)	91	Identities = 130/140 (92%), Positives = 135/140 (95%) gb AAAF98278.1 AF197933_8 (AF197933) beta-hydroxyacyl-ACP dehydratase [Streptococcus pneumoniae] Length = 140
SeqID 1845	SA-471.1	Contig135 (57468-58838 m)	85	Identities = 361/451 (80%), Positives = 405/451 (89%) gb AAAF98279.1 AF197933_9 (AF197933) acetyl-CoA carboxylase biotin carboxylase subunit [Streptococcus pneumoniae] Length = 455

SeqID 1846	SA-472.1	Contig135 (56584-57459 m)	83	Identities = 221/285 (77%), Positives = 248/285 (86%), Gaps = 1/285 (0%) gb AAAF98280.1 AF197933_10 (AF197933) acetyl-CoA carboxylase beta subunit [Streptococcus pneumoniae] Length = 288
SeqID 1847	SA-473.1	Contig135 (55818-56591 m)	83	Identities = 186/254 (73%), Positives = 222/254 (87%) gb AAF98281.1 AF197933_11 (AF197933) acetyl-CoA carboxylase alpha subunit [Streptococcus pneumoniae] Length = 255
SeqID 1848	SA-474.1	Contig135 (54716-55264 m)	45	Identities = 76/142 (53%), Positives = 99/142 (69%) dbj BAA88824.1 (AB016077) sakacin A production response regulator [Streptococcus mutans] Length = 149
SeqID 1849	SA-475.1	Contig135 (53393-54670 p)	74	Identities = 262/425 (61%), Positives = 322/425 (75%), Gaps = 1/425 (0%) sp P37464 SYS_BACSU SERYL-TRNA SYNTHETASE (SERINE--TRNA LIGASE) (SERRS) pir S66043 serine--tRNA ligase (EC 6.1.1.11) - Bacillus subtilis dbj BAA05249.1 (D26185) seryl-tRNA synthetase [Bacillus subtilis] emb CAB11789.1 (Z99104) seryl-tRNA synthetase [Bacillus subtilis] Length = 425
SeqID 1850	SA-477.1	Contig135 (52035-53102 m)	40	Identities = 72/330 (21%), Positives = 143/330 (42%), Gaps = 32/330 (9%) emb CAA07406.1 (AJ006986) transmembrane protein [Streptococcus pneumoniae] Length = 332
SeqID 1851	SA-478.1	Contig135 (51635-51997 p)	75	Identities = 88/112 (78%), Positives = 96/112 (85%) gb AAD46488.1 AF130465_4 (AF130465) unknown [Streptococcus salivarius] Length = 124
SeqID 1852	SA-48.1	Contig137 (2779-2871 m)	No Hits found	
SeqID 1853	SA-480.1	Contig135 (50605-51516 p)	90	Identities = 247/303 (81%), Positives = 276/303 (90%) gb AAD46487.1 AF130465_3 (AF130465) mannose-specific phosphotransferase system component IID [Streptococcus salivarius] Length = 303
SeqID 1854	SA-481.1	Contig135 (49778-50590 p)	84	Identities = 209/271 (77%), Positives = 233/271 (85%), Gaps = 1/271 (0%) gb AAD46486.1 AF130465_2 (AF130465) mannose-specific phosphotransferase system component IIC [Streptococcus salivarius] Length = 271

SeqID 1855	SA-483.1	Contig135 (48735-49745 p)	89	Identities = 287/336 (85%), Positives = 306/336 (90%), Gaps = 6/336 (1%) gb AAD46485.1 AF130465_1 (AF130465) mannose-specific phosphotransferase system component IIAB [Streptococcus salivarius] Length = 330
SeqID 1856	SA-484.1	Contig135 (47620-48432 p)	50	Identities = 89/267 (33%), Positives = 139/267 (51%), Gaps = 3/267 (1%) dbj BAB06625.1 (AP001517) unknown conserved protein [Bacillus halodurans] Length = 269
SeqID 1857	SA-485.1	Contig135 (46947-47531 m)	No Hits found	
SeqID 1858	SA-486.1	Contig135 (46246-46773 m)	44	Identities = 52/189 (27%), Positives = 92/189 (48%), Gaps = 12/189 (6%) pir C71375 conserved hypothetical integral membrane protein TP0033 - syphilis spirochete gb AAC65028.1 (AE001188) conserved hypothetical integral membrane protein [Treponema pallidum] Length = 203
SeqID 1859	SA-487.1	Contig135 (45206-45931 p)	No Hits found	
SeqID 1860	SA-488.1	Contig135 (44722-46143 m)	58	Identities = 193/471 (40%), Positives = 286/471 (59%), Gaps = 42/471 (8%) pir B82096 conserved hypothetical protein VC2278 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF95422.1 (AE004299) conserved hypothetical protein [Vibrio cholerae] Length = 430
SeqID 1861	SA-489.1	Contig135 (44130-44573 m)	53	Identities = 68/149 (45%), Positives = 101/149 (67%), Gaps = 1/149 (0%) dbj BAB04264.1 (AP001508) unknown conserved protein [Bacillus halodurans] Length = 157
SeqID 1862	SA-49.2	Contig137 (2519-2611 p)	No Hits found	
SeqID 1863	SA-490.1	Contig135 (43616-44137 m)	49	Identities = 56/163 (34%), Positives = 94/163 (57%), Gaps = 1/163 (0%) pir D72360 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD35662.1 AE001732_7 (AE001732) conserved hypothetical protein [Thermotoga maritima] Length = 179

SeqID 1864	SA-491.1	Contig135 (42300-43607 m)	40	Identities = 108/324 (33%), Positives = 178/324 (54%), Gaps = 33/324 (10%) sp Q02115 LYTR_BACSU MEMBRANE-BOUND PROTEIN LYTR pir A47679 lyt divergon expression attenuator Lytr - Bacillus subtilis gb AAA22578.1 (M87645) membrane bound protein [Bacillus subtilis] emb CAB15582.1 (Z99122) membrane-bound protein [Bacillus subtilis] Length = 306
SeqID 1865	SA-493.1	Contig135 (41940-42236 p)	No Hits found	Identities = 74/137 (54%), Positives = 95/137 (69%), Gaps = 2/137 (1%) dbj BAB04908.1 (AP001511) Hit-like protein involved in cell-cycle regulation [Bacillus halodurans] Length = 142
SeqID 1866	SA-494.1	Contig135 (41524-41943 p)	65	
SeqID 1867	SA-495.1	Contig135 (40685-41272 m)	No Hits found	
SeqID 1868	SA-497.1	Contig135 (38248-40410 p)	66	Identities = 381/652 (58%), Positives = 485/652 (73%), Gaps = 15/652 (2%) sp O34580 PCRA_BACSU ATP-DEPENDENT DNA HELICASE PCRA pir E69794 ATP-dependent DNA helicase homolog yerF - Bacillus subtilis emb CAA75552.1 (Y15254) PcrA protein [Bacillus subtilis] emb CAB12481.1 (Z99107) similar to ATP-dependent DNA helicase [Bacillus subtilis] Length = 739
SeqID 1869	SA-498.1	Contig135 (37777-38163 p)	53	Identities = 61/87 (70%), Positives = 71/87 (81%) gb AAA88579.1 (M14339) unknown [Streptococcus pneumoniae] Length = 93
SeqID 1870	SA-499.1	Contig135 (36427-37644 p)	62	Identities = 208/423 (49%), Positives = 293/423 (69%), Gaps = 11/423 (2%) sp P41006 PYRP_BACCL URACIL PERMEASE (URACIL TRANSPORTER) pir S38893 uracil transport protein - Bacillus caldolyticus emb CAA53697.1 (X76083) uracil permease [Bacillus caldolyticus] Length = 432
SeqID 1871	SA-5.1	Contig137 (39530-39841 p)	No Hits found	
SeqID 1872	SA-50.1	Contig137 (2617-2736 p)	No Hits found	

SeqID 1873	SA-500.1	Contig135 (34376-35722 m)	69	Identities = 256/443 (57%), Positives = 333/443 (74%), Gaps = 4/443 (0%) sp P44917 Y883_HAEIN HYPOTHETICAL PROTEIN HI0883 pir H64099 probable amino acid transport protein HI0883, sodium-dependent - Haemophilus influenzae (strain Rd KW20) gb AAC22541.1 (U32770) amino acid carrier protein, putative [Haemophilus influenzae Rd] Length = 456
SeqID 1874	SA-501.1	Contig135 (33077-34312 m)	43	Identities = 118/282 (41%), Positives = 181/282 (63%) sp P46348 YEAB_BACSU HYPOTHETICAL 31.8 KD PROTEIN IN GABP-GUAA INTERGENIC REGION (ORFX) pir B69791 cation efflux system membrane protein homolog yeaB - Bacillus subtilis gb AAB62307.1 (U51115) YeaB [Bacillus subtilis] emb CAB12451.1 (Z99107) alternate gene name: ydxT~similar to cation efflux system membrane protein [Bacillus subtilis] Length = 290
SeqID 1875	SA-502.1	Contig135 (32526-32918 m)	49	Identities = 34/110 (30%), Positives = 67/110 (60%), Gaps = 1/110 (0%) pir A71191 hypothetical protein PH1801 - Pyrococcus horikoshii dbj BAA30920.1 (AP000007) 109aa long hypothetical protein [Pyrococcus horikoshii] Length = 109
SeqID 1876	SA-503.1	Contig135 (31850-32545 m)	53	Identities = 80/226 (35%), Positives = 136/226 (59%), Gaps = 1/226 (0%) pir D69983 conserved hypothetical protein ysbB - Bacillus subtilis emb CAA99613.1 (Z75208) hypothetical protein [Bacillus subtilis] emb CAB14850.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 231
SeqID 1877	SA-504.1	Contig135 (31153-31782 m)	42	Identities = 44/194 (22%), Positives = 90/194 (45%), Gaps = 13/194 (6%) emb CAA76857.1 (Y17797) hypothetical protein [Enterococcus faecalis] Length = 247
SeqID 1878	SA-505.1	Contig135 (30605-30751 p)	No Hits found	
SeqID 1879	SA-506.1	Contig135 (30121-30657 m)	No Hits found	
SeqID 1880	SA-507.1	Contig135 (29390-29749 m)	No Hits found	

SeqID 1881	SA-508.2	Contig135 (28209-28997 m)	53	Identities = 124/246 (50%), Positives = 161/246 (65%), Gaps = 22/246 (8%) pir C28551 hypothetical protein 3 - Streptococcus mutans (strain GS-5) (fragment) gb AAA88585.1 (M18954) unknown protein [Streptococcus mutans] Length = 228
SeqID 1882	SA-509.2	Contig135 (73667-74551 p)	66	Identities = 150/285 (52%), Positives = 198/285 (68%), Gaps = 5/285 (1%) gb AAK04462.1 AE006273_8 (AE006273) pseudouridine synthase [Lactococcus lactis subsp. lactis] Length = 293
SeqID 1883	SA-51.1	Contig137 (2184-2513 p)	No Hits found	
SeqID 1884	SA-510.1	Contig135 (72831-73670 p)	61	Identities = 120/267 (44%), Positives = 174/267 (64%), Gaps = 3/267 (1%) sp O31612 YJBN_BACSU HYPOTHETICAL 30.0 KDA PROTEIN IN MECA-TENA INTERGENIC REGION pir F69844 conserved hypothetical protein yjBN - Bacillus subtilis emb CAB13018.1 (Z99110) similar to hypothetical proteins [Bacillus subtilis] Length = 266
SeqID 1885	SA-511.1	Contig135 (72188-72859 p)	69	Identities = 115/208 (55%), Positives = 159/208 (76%), Gaps = 3/208 (1%) dbj BAB06568.1 (AP001516) GTP pyrophosphokinase [Bacillus halodurans] Length = 211
SeqID 1886	SA-512.1	Contig135 (71506-72078 m)	46	Identities = 63/184 (34%), Positives = 99/184 (53%), Gaps = 10/184 (5%) pir C69844 hypothetical protein yjbK - Bacillus subtilis emb CAB13015.1 (Z99110) yjbK [Bacillus subtilis] Length = 190

SeqID 1887	SA-513.1	Contig135 (70355-71329 m)	69	<p>Identities = 166/319 (52%), Positives = 231/319 (72%), Gaps = 4/319 (1%) sp P14193 KPRS_BACSU_RIBOSE-PHOSPHATE PYROPHOSPHOKINASE (PHOSPHORIBOSYL PYROPHOSPHATE SYNTHETASE) (PRPP SYNTHETASE) pir KBSRS ribose-phosphate pyrophosphokinase (EC 2.7.6.1) prs - Bacillus subtilis pdb 1DKR B Chain B, Crystal Structures Of Bacillus Subtilis Phosphoribosylpyrophosphate Synthetase: Molecular Basis Of Allosteric Inhibition And Activation. pdb 1DKR A Chain A, Crystal Structures Of Bacillus Subtilis Phosphoribosylpyrophosphate Synthetase: Molecular Basis Of Allosteric Inhibition And Activation. pdb 1DKU A Chain A, Crystal Structures Of Bacillus Subtilis Phosphoribosylpyrophosphate Synthetase: Molecular Basis Of Allosteric Inhibition And Activation. pdb 1DKU B Chain B, Crystal Structures Of Bacillus Subtilis Phosphoribosylpyrophosphate Synthetase: Molecular Basis Of Allosteric Inhibition And Activation. emb CAA34523.1 (X16518) PRPP synthetase (AA</p>
SeqID 1888	SA-514.1	Contig135 (69920-70354 m)	55	<p>Identities = 63/118 (53%), Positives = 81/118 (68%), Gaps = 1/118 (0%) gb AAG01802.1 AF276772_1 (AF276772) cysteine desulfurase Nifs [Methanosarcina thermophila] Length = 404</p>
SeqID 1889	SA-515.1	Contig135 (69318-69647 m)	No Hits found	
SeqID 1890	SA-517.1	Contig135 (67303-68655 p)	69	<p>Identities = 223/448 (49%), Positives = 313/448 (69%) sp P94417 AK3_BACSU_PROBABLE ASPARTOKINASE (ASPARTATE KINASE) pir A69763 homoserine dehydrogenase homolog yciM - Bacillus subtilis dbj BAA09011.1 (D50453) homologue of aspartokinase 2 alpha and beta subunits LysC of B. subtilis [Bacillus subtilis] emb CAB12187.1 (Z99106) similar to homoserine dehydrogenase [Bacillus subtilis] Length = 454</p>

SeqID 1891	SA-518.1	Contig135 (66559-67209 m)	42	Identities = 59/190 (31%), Positives = 93/190 (48%), Gaps = 10/190 (5%) pir JAB82498 CbbY-family-protein-VCA0102 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF96016.1 (AE004353) CbbY family protein [Vibrio cholerae] Length = 219
SeqID 1892	SA-519.3	Contig135 (65631-66422 m)	64	Identities = 139/248 (56%), Positives = 185/248 (74%) dbj BAB16033.1 (AB030809) Pseudomonas putida enoyl-CoA hydratase II homologue [Streptococcus pyogenes] Length = 248
SeqID 1893	SA-52.1	Contig137 (1679-2173 p)	No Hits found	
SeqID 1894	SA-520.2	Contig116 (395-688 m)	42	Identities = 24/86 (27%), Positives = 43/86 (49%), Gaps = 6/86 (6%) gb AAD33114.1 (AF094574_4 (AF094574) negative regulator of translation [Haemophilus influenzae] Length = 98
SeqID 1895	SA-523.2	Contig116 (1354-4734 m)	73	Identities = 767/1103 (69%), Positives = 836/1103 (75%), Gaps = 58/1103 (5%) gb AAD39085.1 (AF091393_1 (AF091393) surface protein R28 [Streptococcus pyogenes] Length = 1260
SeqID 1896	SA-524.1	Contig116 (4977-6170 p)	45	Identities = 187/187 (100%), Positives = 187/187 (100%) gb AAG09971.1 (AF248037_6 (AF248037) unknown [Streptococcus agalactiae] Length = 189
SeqID 1897	SA-526.1	Contig116 (6365-6889 m)	97	Identities = 174/174 (100%), Positives = 174/174 (100%) gb AAG09969.1 (AF248037_4 (AF248037) unknown [Streptococcus agalactiae] Length = 174
SeqID 1898	SA-527.2	Contig116 (7021-7599 p)	90	Identities = 177/198 (89%), Positives = 180/198 (90%), Gaps = 11/198 (5%) gb AAG09968.1 (AF248037_3 (AF248037) cation efflux system protein [Streptococcus agalactiae] Length = 287
SeqID 1899	SA-528.2	Contig116 (7971-8354 p)	92	Identities = 125/128 (97%), Positives = 128/128 (99%) gb AAG09967.1 (AF248037_2 (AF248037) aldose reductase [Streptococcus agalactiae] Length = 280
SeqID 1900	SA-529.1	Contig116 (8355-8825 p)	96	Identities = 152/156 (97%), Positives = 152/156 (97%), Gaps = 4/156 (2%) gb AAG09967.1 (AF248037_2 (AF248037) aldose reductase [Streptococcus agalactiae] Length = 280

SeqID 1901	SA-53.1	Contig137 (1454-1642 p)	No Hits found	
SeqID 1902	SA-530.1	Contig116 (8935-9390 p)	58	Identities = 90/90 (100%), Positives = 90/90 (100%) gb AAG09966.1 AF248037_1 (AF248037) alcohol dehydrogenase [Streptococcus agalactiae] Length = 96
SeqID 1903	SA-531.1	Contig116 (9315-9971 p)	65	Identities = 114/231 (49%), Positives = 161/231 (69%), Gaps = 6/231 (2%) gb AAG20655.1 (AE005134) alcohol dehydrogenase; Adh2 [Halobacterium sp. NRC-1] Length = 347
SeqID 1904	SA-532.1	Contig116 (9987-10376 p)	64	Identities = 53/123 (43%), Positives = 84/123 (68%) pir B69970 transcription regulator MerR family homolog yraB - Bacillus subtilis emb CAA63468.1 (X92868) mercuric resistance operon regulatory protein [Bacillus subtilis] emb CAB14642.1 (Z99117) similar to transcriptional regulator (MerR family) [Bacillus subtilis] Length = 140
SeqID 1905	SA-533.1	Contig116 (10386-10781 p)	54	Identities = 55/131 (41%), Positives = 73/131 (54%), Gaps = 8/131 (6%) pir B72308 hypothetical protein - Thermotoga maritima (strain MSB8) gb AAD36075.1 AE001762_2 (AE001762) hypothetical protein [Thermotoga maritima] Length = 135
SeqID 1906	SA-534.1	Contig116 (10805-11122 p)	48	Identities = 33/93 (35%), Positives = 59/93 (62%), Gaps = 1/93 (1%) pir T29425 4-carboxymuconolactone decarboxylase homolog - Streptomyces coelicolor emb CAA20070.1 (AL031155) 3-oxoadipate enol-lactone hydrolase/4- carboxymuconolactone decarboxylase [Streptomyces coelicolor A3(2)] Length = 449
SeqID 1907	SA-535.1	Contig116 (11209-11493 p)	No Hits found	

SeqID 1908	SA-536.1	Contig116 (11732-12223 p)	42	Identities = 39/149 (26%), Positives = 71/149 (47%), Gaps = 4/149 (2%) pir H83035-probable transcription regulator PA4878 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG08263.1 AE004901_5 (AE004901) probable transcriptional regulator-[Pseudomonas aeruginosa] Length = 270
SeqID 1909	SA-537.1	Contig116 (12267-12656 m)	52	Identities = 42/130 (32%), Positives = 71/130 (54%), Gaps = 3/130 (2%) pir G69153 conserved hypothetical protein MTH413 - Methanobacterium thermoautotrophicum (strain Delta H) gb AAB84919.1 (AE000825) conserved protein [Methanobacterium thermoautotrophicum] Length = 130
SeqID 1910	SA-538.1	Contig116 (12669-15836 m)	22	Identities = 141/566 (24%), Positives = 242/566 (41%), Gaps = 52/566 (9%) pir T31094 surface antigen BspA - Bacteroides forsythus gb AAC82625.1 (AF054892) surface antigen BspA [Bacteroides forsythus] Length = 1081
SeqID 1911	SA-54.1	Contig137 (321-1361 p)	53	Identities = 102/413 (24%), Positives = 189/413 (45%), Gaps = 76/413 (18%) sp P55340 ECSB_BACSU PROTEIN ECSB pir G69619 ABC transporter (membrane protein) ecsB - Bacillus subtilis emb CAA61075.1 (X87807) hypothetical EcsB protein [Bacillus subtilis] emb CAA74408.1 (Y14077) Hypothetical protein [Bacillus subtilis] emb CAB12845.1 (Z99109) ABC transporter (membrane protein) [Bacillus subtilis] Length = 408
SeqID 1912	SA-540.1	Contig116 (15899-18067 m)	79	Identities = 475/727 (65%), Positives = 585/727 (80%), Gaps = 19/727 (2%) sp P78027 IR1_MYCPN RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE ALPHA CHAIN (RIBONUCLEOTIDE REDUCTASE) pir S73838 ribonucleoside-diphosphate reductase (EC 1.17.4.1) large chain nrdE - Mycoplasma pneumoniae (strain ATCC 29342) gb AAB96160.1 (AE000050) ribonucleoside-diphosphate reductase alpha chain~MPN324(new), 513(Himmelreich et al., 1996) [Mycoplasma pneumoniae] Length = 721

SeqID 1913	SA-541.1	Contig116 (18069-18470 m)	76	Identities = 76/127 (59%), Positives = 105/127 (81%), Gaps = 1/127 (0%) sp P47472 NRDI_MYCGE_NRD1_PROTEIN pir D64225 hypothetical protein MG230 - Mycoplasma genitalium gb AAC71451.1 (U39702) nrd1 protein (nrd1) [Mycoplasma genitalium] Length = 153
SeqID 1914	SA-542.1	Contig116 (18483-19493 m)	87	Identities = 259/335 (77%), Positives = 299/335 (88%) sp P75461 RIR2_MYCPN_RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE BETA CHAIN (RIBONUCLEOTIDE REDUCTASE) pir S73840 ribonucleotide reductase 2 - Mycoplasma pneumoniae (strain ATCC 29342) gb AAB96162.1 (AE00050) ribonucleoside-diphosphate reductase beta chain [Mycoplasma pneumoniae] Length = 339
SeqID 1915	SA-544.1	Contig116 (20046-20954 m)	56	Identities = 105/309 (33%), Positives = 173/309 (55%), Gaps = 21/309 (6%) pir T00087 rhamnosyltransferase - Streptococcus mutans dbj BAA32090.1 (AB010970) rhamnosyltransferase [Streptococcus mutans] Length = 311
SeqID 1916	SA-545.2	Contig116 (21204-21404 p)	No Hits found	
SeqID 1917	SA-546.1	Contig116 (24213-24722 p)	No Hits found	
SeqID 1918	SA-547.1	Contig116 (21164-24883 m)	53	Identities = 461/1194 (38%), Positives = 666/1194 (55%), Gaps = 74/1194 (6%) gb AAB17762.1 (U56908) SCPB [Streptococcus agalactiae] Length = 1150
SeqID 1919	SA-548.1	Contig116 (25053-25481 m)	60	Identities = 65/143 (45%), Positives = 93/143 (64%), Gaps = 5/143 (3%) pir H83632 conserved hypothetical protein PA0115 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG03505.1 AE004449_14 (AE004449) conserved hypothetical protein [Pseudomonas aeruginosa] Length = 150
SeqID 1920	SA-549.1	Contig116 (25515-26279 m)	32	Identities = 48/180 (26%), Positives = 84/180 (46%), Gaps = 4/180 (2%) pir F82497 uridine phosphorylase VCA0134 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF96047.1 (AE004354) uridine phosphorylase [Vibrio cholerae] Length = 243

SeqID 1921	SA-55.1	Contig137 (2-319 p)	77	Identities = 64/104 (61%), Positives = 82/104 (78%) sp P55339 ECSA_BACSU_ABC-TYPE_TRANSPORTER_ATP-BINDING_PROTEIN_ECSA_pir F69619 ABC transporter (ATP-binding protein) ecsA - Bacillus subtilis emb CAA61074.1 (X87807) putative ATP-binding protein of ABC-type [Bacillus subtilis] emb CAA74409.1 (Y14077) Hypothetical protein [Bacillus subtilis] emb CAB12844.1 (Z99109) ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 247
SeqID 1922	SA-550.1	Contig116 (26546-27901 m)	64	Identities = 216/448 (48%), Positives = 297/448 (66%), Gaps = 4/448 (0%) pir F69806 RNA methyltransferase homolog yfjO - Bacillus subtilis emb CAB12631.1 (Z99108) similar to RNA methyltransferase [Bacillus subtilis] dbj BAA24300.1 (D78509) YfjO [Bacillus subtilis] Length = 466
SeqID 1923	SA-551.1	Contig116 (28099-28776 p)	50	Identities = 73/263 (27%), Positives = 140/263 (52%), Gaps = 9/263 (3%) dbj BAB04643.1 (AP001510) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 266
SeqID 1924	SA-552.1	Contig116 (28860-29393 p)	68	Identities = 96/175 (54%), Positives = 122/175 (68%) dbj BAB04659.1 (AP001510) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 175
SeqID 1925	SA-554.1	Contig121 (35263-36249 p)	60	Identities = 142/331 (42%), Positives = 204/331 (60%), Gaps = 2/331 (0%) gb AAF61315.1 (U96166) unknown [Streptococcus cristatus] Length = 442
SeqID 1926	SA-555.1	Contig121 (33762-35270 p)	19	Identities = 66/194 (34%), Positives = 98/194 (50%), Gaps = 9/194 (4%) dbj BAA94320.1 (AB033763) hypothetical protein [Staphylococcus aureus] Length = 255

SeqID 1927	SA-557.1	Contig121 (31361-33748 p)	63	Identities = 336/794 (42%), Positives = 507/794 (63%), Gaps = 29/794 (3%) sp P28366 SEGA_BAGSU PREPROTEIN TRANSLOCASE SECA SUBUNIT pir JQ0647 preprotein translocase secA - Bacillus subtilis dbj BAA01122.1 (D10279) secA protein [Bacillus subtilis] gb AAC44957.1 (U56901) involved in protein export [Bacillus subtilis] emb CAB15547.1 (Z99122) translocase binding subunit (ATPase) [Bacillus subtilis] Length = 841
SeqID 1928	SA-558.1	Contig121 (30364-31374 p)	No Hits found	
SeqID 1929	SA-559.1	Contig121 (28826-30385 p)	No Hits found	
SeqID 1930	SA-56.1	Contig131 (56872-57270 p)	No Hits found	
SeqID 1931	SA-560.1	Contig121 (27275-28819 p)	No Hits found	
SeqID 1932	SA-561.1	Contig121 (26046-27275 p)	50	Identities = 105/422 (24%), Positives = 213/422 (49%), Gaps = 49/422 (11%) pir C82917 preprotein translocase UU250 [imported] - Ureaplasma urealyticum gb AAF30659.1 AE002122_28 (AE002122) preprotein translocase [Ureaplasma urealyticum] Length = 471
SeqID 1933	SA-562.1	Contig121 (24726-25922 p)	10	Identities = 30/78 (38%), Positives = 42/78 (53%) gb AAF61315.1 (U96166) unknown [Streptococcus cristatus] Length = 442
SeqID 1934	SA-563.1	Contig121 (23790-24665 p)	43	Identities = 88/228 (38%), Positives = 136/228 (59%), Gaps = 10/228 (4%) gb AAC44016.1 (U40830) EpsI [Streptococcus thermophilus] prf J2209356K epsi gene [Streptococcus thermophilus] Length = 324
SeqID 1935	SA-565.1	Contig121 (22613-23797 p)	35	Identities = 91/256 (35%), Positives = 146/256 (56%), Gaps = 8/256 (3%) gb AAF28363.1 AF224467_2 (AF224467) putative glycosyl transferase [Haemophilus ducreyi] Length = 269
SeqID 1936	SA-566.1	Contig121 (21382-22623 p)	37	Identities = 103/259 (39%), Positives = 156/259 (59%), Gaps = 3/259 (1%) gb AAF28363.1 AF224467_2 (AF224467) putative glycosyl transferase [Haemophilus ducreyi] Length = 269

SeqID 1937	SA-567.1	Contig121 (20180-21385 p)	38	Identities = 94/263 (35%), Positives = 158/263 (59%), Gaps = 4/263 (1%) gb AAAF28363.1 AF224467_2 (AF224467) putative glycosyl transferase [Haemophilus ducreyi] Length = 269
SeqID 1938	SA-568.1	Contig121 (19164-20171 p)	33	Identities = 68/286 (23%), Positives = 122/286 (41%), Gaps = 18/286 (6%) sp P37749 YEFG_ECOLI HYPOTHETICAL 37.8 KD PROTEIN IN GND-RFC INTERGENIC REGION (GALF TRANSFERASE) pir 69645 probable nucleotide sugar synthetase - Escherichia coli gb AAB88405.1 (U09876) putative Galf transferase [Escherichia coli] gb AAC31634.1 (U03041) nucleotide sugar synthetase [Escherichia coli] dbj BAA15876.1 (D90841) ORF_ID:0351 10~similar to [SwissProt Accession Number P37749] [Escherichia coli] dbj BAA15888.1 (D90842) ORF_ID:0351 10; similar to [SwissProt Accession Number P37749] [Escherichia coli] gb AAC75095.1 (AE000294) putative Galf transferase [Escherichia coli K12] Length = 330
SeqID 1939	SA-569.1	Contig121 (16766-18616 m)	42	Identities = 180/657 (27%), Positives = 283/657 (42%), Gaps = 31/657 (4%) emb CAB65343.1 (AJ007010) liver stage antigen-3 [Plasmodium falciparum] Length = 1786
SeqID 1940	SA-57.1	Contig131 (56480-56863 p)	No Hits found	
SeqID 1941	SA-570.1	Contig121 (15272-15448 m)	No Hits found	
SeqID 1942	SA-571.1	Contig121 (14901-18833 p)	61	Identities = 668/1314 (50%), Positives = 812/1314 (60%), Gaps = 60/1314 (4%) dbj BAA97453.1 (AB029393) streptococcal hemagglutinin [Streptococcus gordonii] Length = 2178
SeqID 1943	SA-572.1	Contig121 (13022-14518 m)	59	Identities = 201/493 (40%), Positives = 302/493 (60%), Gaps = 5/493 (1%) pir S43609 rofA protein - Streptococcus pyogenes Length = 497

SeqID 1944	SA-573.1	Contig121 (10934-12925 p)	94	Identities = 579/667 (86%), Positives = 635/667 (94%), Gaps = 1/667 (0%) sp Q54986 UVRB_STRPN EXCINUCLEASE ABC SUBUNIT B pir A42385 excinuclease ABC chain B - Streptococcus pneumoniae plasmid pSB470 gb AAA27020.1 (M80215)-uvs402-protein [Streptococcus pneumoniae] Length = 668
SeqID 1945	SA-574.1	Contig121 (10034-10873 p)	29	Identities = 58/190 (30%), Positives = 96/190 (50%), Gaps = 11/190 (5%) pir T34651 probable transmembrane protein - Streptomyces coelicolor emb CAA22372.1 (AL034446) putative transmembrane protein [Streptomyces coelicolor A3(2)] Length = 307
SeqID 1946	SA-575.1	Contig121 (7587-9770 m)	21	Identities = 109/195 (55%), Positives = 156/195 (79%), Gaps = 4/195 (2%) gb AAF16724.1 AF141644_1 (AF141644) putative integral membrane protein [Lactococcus lactis] Length = 191
SeqID 1947	SA-576.1	Contig121 (6847-7587 m)	77	Identities = 147/240 (61%), Positives = 192/240 (79%) ref NP_069514.1 glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus] pir H69334 glutamine transport protein glnQ - Archaeoglobus fulgidus gb AAB90561.1 (AE001058) glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus] Length = 242
SeqID 1948	SA-577.1	Contig121 (6348-6707 m)	No Hits found	
SeqID 1949	SA-578.1	Contig121 (6098-6256 p)	No Hits found	
SeqID 1950	SA-579.1	Contig121 (4759-6072 p)	77	Identities = 297/435 (68%), Positives = 345/435 (79%), Gaps = 7/435 (1%) sp P20964 OBG_BACSU SPO0B-ASSOCIATED GTP-BINDING PROTEIN pir B32804 GTP-binding protein obg - Bacillus subtilis gb AAA22505.1 (M24537) GTP-binding protein [Bacillus subtilis] emb CAB14752.1 (Z99118) GTPase activity [Bacillus subtilis] Length = 428

SeqID 1951	SA-58.1	Contig131 (55941-56345 p)	22	Identities = 25/79 (31%), Positives = 31/79 (38%) sp P04929 HRPX_PLALO HISTIDINE-RICH-GLYCOPROTEIN PRECURSOR pir KGZQHL histidine-rich glycoprotein precursor Plasmodium lophurae emb CAA25698.1 (X01469) histidine-rich protein [Plasmodium lophurae] prf 1101401A.protein.His rich [Plasmodium sp.] Length = 351
SeqID 1952	SA-582.1	Contig121 (3151-4392 m)	73	Identities = 247/413 (59%), Positives = 314/413 (75%) sp Q9X4A7 PEPS_STRTR AMINOPEPTIDASE PEPS gb AAD28348.1 AF102860_2 (AF102860) aminopeptidase Peps [Streptococcus thermophilus] Length = 413
SeqID 1953	SA-583.1	Contig121 (2613-2852 m)	No Hits found	
SeqID 1954	SA-584.1	Contig121 (2539-3117 p)	34	Identities = 41/152 (26%), Positives = 75/152 (48%), Gaps = 4/152 (2%) emb CAB88235.1 (AL353012) hypothetical serine- rich repeat protein [Schizosaccharomyces pombe] Length = 451
SeqID 1955	SA-585.2	Contig121 (387-2429 p)	38	Identities = 173/492 (35%), Positives = 270/492 (54%), Gaps = 38/492 (7%) gb AAK04264.1 AE006254_5 (AE006254) amidase [Lactococcus lactis subsp. lactis] Length = 499
SeqID 1956	SA-586.2	Contig121 (3-302 p)	57	Identities = 39/89 (43%), Positives = 57/89 (63%), Gaps = 4/89 (4%) dbj BAB06992.1 (AP001518) 16S pseudouridylate synthase [Bacillus halodurans] Length = 238
SeqID 1957	SA-589.2	Contig122 (32148-33029 m)	86	Identities = 230/291 (79%), Positives = 257/291 (88%) sp Q07211 SCRK_STRMU FRUCTOKINASE dbj BAA02467.1 (D13175) fructokinase [Streptococcus mutans] Length = 293
SeqID 1958	SA-59.1	Contig131 (55843-55980 m)	No Hits found	
SeqID 1959	SA-590.1	Contig122 (31083-32030 m)	82	Identities = 232/312 (74%), Positives = 262/312 (83%) sp Q59935 MANA_STRMU MANNOSE-6-PHOSPHATE ISOMERASE (PHOSPHOMANNOSE ISOMERASE) (PMI) (PHOSPHOHEXOMUTASE) dbj BAA04021.1 (D16594) Mannosephosphate Isomerase [Streptococcus mutans] Length = 316

SeqID 1960	SA-591.1	Contig122 (28446-30974 m)	75	Identities = 503/843 (59%), Positives = 643/843 (75%), Gaps = 16/843 (4%) sp P47847 SECA_LISMO PREPROTEIN - TRANSLOCASE SECA SUBUNIT gb AAA50286.1 (L32090) secA [Listeria monocytogenes] Length = 836
SeqID 1961	SA-592.1	Contig122 (27313-28320 m)	54	Identities = 122/348 (35%), Positives = 188/348 (53%), Gaps = 32/348 (9%) pir A81791 2-dehydro-3-deoxyphosphoheptonate aldolase (EC 4.1.2.15) NMA2180 [imported] - Neisseria meningitidis (group A strain Z2491) emb CAB85392.1 (AL162758) phospho-2-dehydr-3-deoxyheptonate aldolase [Neisseria meningitidis] Length = 351
SeqID 1962	SA-593.1	Contig122 (26929-27288 m)	73	Identities = 61/120 (50%), Positives = 91/120 (75%), Gaps = 1/120 (0%) gb AAG22706.1 (AF276617) acyl carrier protein synthase; Acps [Streptococcus pneumoniae] Length = 120
SeqID 1963	SA-594.1	Contig122 (25832-26932 m)	72	Identities = 227/366 (62%), Positives = 270/366 (73%) gb AAD51027.1 AF171873_1 (AF171873) alanine racemase [Streptococcus pneumoniae] Length = 367
SeqID 1964	SA-597.1	Contig122 (24201-25739 m)	66	Identities = 248/530 (46%), Positives = 341/530 (63%), Gaps = 18/530 (3%) gb AAB52379.1 (U31811) immunogenic secreted protein precursor [Streptococcus pyogenes] Length = 534
SeqID 1965	SA-598.1	Contig122 (22106-24121 m)	84	Identities = 483/671 (71%), Positives = 568/671 (83%) sp Q54900 RECG_STRPN ATP-DEPENDENT DNA HELICASE RECG pir S71016 helicase recG homolog - Streptococcus pneumoniae emb CAA90280.1 (Z49988) MmsA [Streptococcus pneumoniae] prf 2209420A mmsA gene [Streptococcus pneumoniae] Length = 671
SeqID 1966	SA-599.1	Contig122 (20901-21821 m)	72	Identities = 173/300 (57%), Positives = 224/300 (74%) dbj BAB07646.1 (AP001520) aryl-alcohol dehydrogenase [Bacillus halodurans] Length = 305

SeqID 1967	SA-6.1	Contig137 (39066-39299 p)	42	Identities = 31/81 (38%), Positives = 43/81 (52%), Gaps = 18/81 (22%) sp P43813 BNLJ_HAEIN_DNA_LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE [NAD+]) pir D64182 DNA_ligase (NAD+) (EC 6.5.1.2) - Haemophilus influenzae (strain Rd KW20) gb AAC22753.1 (U32789) DNA_ligase (lig) [Haemophilus influenzae Rd] Length = 679
SeqID 1968	SA-60.1	Contig131 (55778-56308 p)	31	Identities = 42/95 (44%), Positives = 56/95 (58%), Gaps = 4/95 (4%) ref XP_001738.1 megakaryocyte stimulating factor [Homo sapiens] Length = 1385
SeqID 1969	SA-600.1	Contig122 (19925-20803 m)	56	Identities = 114/279 (40%), Positives = 171/279 (60%), Gaps = 3/279 (1%) sp P28244 YDIB_ECOLI_HYPOTHETICAL_31.2_KDA PROTEIN IN LPP-AROD INTERGENIC REGION pir D64927 probable shikimate 5-dehydrogenase (EC 1.1.1.25) ydiB - Escherichia coli dbj BAA15449.1 (D90811) Shikimate 5-dehydrogenase (EC 1.1.1.25) [Escherichia coli] gb AAC74762.1 (AE000264) putative oxidoreductase [Escherichia coli K12] Length = 288
SeqID 1970	SA-601.1	Contig122 (18928-19890 p)	64	Identities = 158/319 (49%), Positives = 214/319 (66%), Gaps = 4/319 (1%) dbj BAB05343.1 (AP001512) L-asparaginase [Bacillus halodurans] Length = 322
SeqID 1971	SA-602.1	Contig122 (17477-18859 m)	30	Identities = 89/281 (31%), Positives = 141/281 (49%), Gaps = 31/281 (11%) pir C69862 conserved hypothetical protein ykrA - Bacillus subtilis emb CAB13328.1 (Z99111) similar to hypothetical proteins [Bacillus subtilis] gb AAC24929.1 (AF012285) unknown [Bacillus subtilis] Length = 257
SeqID 1972	SA-603.2	Contig122 (16969-17421 p)	55	Identities = 62/141 (43%), Positives = 93/141 (64%) dbj BAB06903.1 (AP001518) BH3184~unknown conserved protein [Bacillus halodurans] Length = 147

SeqID 1973	SA-604.2	Contig122 (15490-16701 m)	79	Identities = 250/404 (61%), Positives = 328/404 (80%), Gaps = 1/404 (0%) sp P71348 YFBQ_HAEIN PROBABLE AMINOTRANSFERASE HI0286 gb AAC21948.1 (U32714) aminotransferase [Haemophilus influenzae Rd] Length = 404
SeqID 1974	SA-605.1	Contig122 (14579-15364 m)	62	Identities = 129/257 (50%), Positives = 181/257 (70%), Gaps = 3/257 (1%) db BAB06181.1 (AP001515) transcriptional pleiotropic repressor [Bacillus halodurans] Length = 259
SeqID 1975	SA-606.1	Contig122 (13964-14512 m)	69	Identities = 101/183 (55%), Positives = 133/183 (72%) pir C70008 pyrazinamidase/nicotinamidase homolog yueJ - Bacillus subtilis emb CAB15164.1 (Z99120) similar to pyrazinamidase/nicotinamidase [Bacillus subtilis] Length = 183
SeqID 1976	SA-607.1	Contig122 (12952-13917 p)	51	Identities = 151/321 (47%), Positives = 196/321 (61%), Gaps = 36/321 (11%) pir B75610 probable 3-hydroxyacyl-CoA dehydrogenase - Deinococcus radiodurans (strain R1) gb AAF12219.1 AE001862_45 (AE001862) 3-hydroxyacyl-CoA dehydrogenase, putative [Deinococcus radiodurans] Length = 347
SeqID 1977	SA-608.1	Contig122 (12134-12646 m)	64	Identities = 88/210 (41%), Positives = 141/210 (66%), Gaps = 3/210 (1%) pir A69969 conserved hypothetical protein yqzB - Bacillus subtilis emb CAB14454.1 (Z99116) similar to hypothetical proteins [Bacillus subtilis] emb CAB14467.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 212
SeqID 1978	SA-609.1	Contig122 (11293-12123 m)	65	Identities = 126/264 (47%), Positives = 186/264 (69%), Gaps = 1/264 (0%) db BAB05092.1 (AP001511) unknown conserved protein [Bacillus halodurans] Length = 270
SeqID 1979	SA-611.1	Contig122 (8635-11280 m)	73	Identities = 495/870 (56%), Positives = 648/870 (73%), Gaps = 4/870 (0%) sp P22983 PODK_CLOSYPYRUVATE,PHOSPHATE DIKINASE (PYRUVATE,ORTHOPHOSPHATE DIKINASE) Length = 874

SeqID 1980	SA-612.1	Contig122 (8195-8497 m)	61	<p>Identities = 42/96 (43%), Positives = 64/96 (65%) dbj BAB04384.1 (AP001509) glutamyl-tRNA (Gln) amidotransferase subunit C [Bacillus halodurans] Length = 96</p>
SeqID 1981	SA-613.1	Contig122 (6729-8195 m)	73	<p>Identities = 285/485 (58%), Positives = 365/485 (74%), Gaps = 2/485 (0%) sp O06491 GATA_BACSU GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (GLU-ADT SUBUNIT A) pir B69795 glutamyl-tRNA(Gln) amidotransferase (EC 2.6.-.-) chain A [validated] - Bacillus subtilis emb CAB12488.1 (Z99107) alternate gene name: yedB~similar to amidase [Bacillus subtilis] Length = 485</p>
SeqID 1982	SA-614.1	Contig122 (5287-6729 m)	74	<p>Identities = 309/476 (64%), Positives = 361/476 (74%), Gaps = 1/476 (0%) sp Q9Z9X0 GATB_BACHD GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT B (GLU-ADT SUBUNIT B) pir T44293 hypothetical protein yerN [imported] - Bacillus halodurans dbj BAA75312.1 (AB011836) similar to B.subtilis yerN gene(87 identity) [Bacillus halodurans] dbj BAB04386.1 (AP001509) glutamyl-tRNA (Gln) amidotransferase subunit B [Bacillus halodurans] Length = 476</p>
SeqID 1983	SA-615.1	Contig122 (4243-5157 m)	54	<p>Identities = 102/292 (34%), Positives = 169/292 (56%), Gaps = 3/292 (1%) pir T30575 hypothetical protein - Bacillus megaterium (fragment) emb CAA04271.1 (AJ000733) hypothetical protein [Bacillus megaterium] Length = 296</p>
SeqID 1984	SA-616.1	Contig122 (3599-4126 m)	42	<p>Identities = 52/153 (33%), Positives = 88/153 (56%) sp P54452 YQEG_BACSU HYPOTHETICAL 20.1 KD PROTEIN IN NUCB-AROD INTERGENIC REGION pir C69951 conserved hypothetical protein yqeG - Bacillus subtilis dbj BAA12443.1 (D84432) YqeG [Bacillus subtilis] emb CAB14510.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 172</p>

SeqID 1985	SA-617.1	Contig122 (2481-3599 m)	73	<p>Identities = 219/373 (58%), Positives = 280/373 (74%), Gaps = 8/373 (2%) sp P54453 YQEH_BACSU-HYPOTHETICAL 41.0 KD PROTEIN IN NUCB-AROD INTERGENIC REGION pir D69951 conserved hypothetical protein yqeH - Bacillus subtilis dbj BAA12444.1 (D84432) YqeH [Bacillus subtilis] emb CAB14509.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 366</p>
SeqID 1986	SA-618.1	Contig122 (2071-2388 m)	56	<p>Identities = 49/97 (50%), Positives = 67/97 (68%), Gaps = 2/97 (2%) sp P54454 YQEI_BACSU HYPOTHETICAL 10.8 KD PROTEIN IN AROD-COMER INTERGENIC REGION pir E69951 dihydrodipicolinate reductase homolog yqeI - Bacillus subtilis dbj BAA12446.1 (D84432) YqeI [Bacillus subtilis] emb CAB14507.1 (Z99117) similar to dihydrodipicolinate reductase [Bacillus subtilis] Length = 96</p>
SeqID 1987	SA-62.1	Contig131 (55074-55907 m)	44	<p>Identities = 77/263 (29%), Positives = 126/263 (47%), Gaps = 34/263 (12%) gb AAC48867.1 (U40453) integrase [Streptococcus pyogenes phage T12] Length = 362</p>
SeqID 1988	SA-620.1	Contig122 (1309-1941 m)	59	<p>Identities = 85/187 (45%), Positives = 134/187 (71%) sp P54455 NADD_BACSU PROBABLE NICOTINATE-NUCLEOTIDE ADENYLYLTRANSFERASE (DEAMIDO-NAD(+)) PYROPHOSPHORYLASE (DEAMIDO-NAD(+)) DIPHOSPHORYLASE (NICOTINATE MONONUCLEOTIDE ADENYLYLTRANSFERASE) (NAMN ADENYLYLTRANSFERASE) pir F69951 conserved hypothetical protein yqeJ - Bacillus subtilis dbj BAA12447.1 (D84432) YqeJ [Bacillus subtilis] emb CAB14506.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 189</p>

SeqID 1989	SA-621.1	Contig122 (725-1312 m)	55	<p>Identities = 79/180 (43%), Positives = 116/180 (63%) sp P54456 YQEK_BACSU HYPOTHETICAL 21.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION pir G69951 conserved hypothetical protein yqk - Bacillus subtilis dbj BAA12448.1 (D84432) Yqk [Bacillus subtilis] emb CAB14505.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] Length = 186</p>
SeqID 1990	SA-622.1	Contig122 (140-682 m)	42	<p>Identities = 46/175 (26%), Positives = 81/175 (46%), Gaps = 12/175 (6%) gb AAG19496.1 (AE005041) Vng1100c [Halobacterium sp. NRC-1] Length = 183</p>
SeqID 1991	SA-623.2	Contig129 (30428-31114 m)	52	<p>Identities = 80/214 (37%), Positives = 122/214 (56%), Gaps = 3/214 (1%) gb AAC95438.1 (AF068901) unknown [Streptococcus pneumoniae] Length = 234</p>
SeqID 1992	SA-624.1	Contig129 (31243-32463 m)	74	<p>Identities = 220/398 (55%), Positives = 306/398 (76%), Gaps = 3/398 (0%) gb AAF36228.1 (AF168363_4 (AF168363) oxalate:formate antiporter [Lactococcus lactis] Length = 421</p>
SeqID 1993	SA-625.1	Contig129 (32650-34017 m)	76	<p>Identities = 313/453 (69%), Positives = 375/453 (82%) gb AAC95436.1 (AF068901) D-Ala-D-Ala adding enzyme [Streptococcus pneumoniae] Length = 457</p>
SeqID 1994	SA-626.1	Contig129 (34164-35210 m)	81	<p>Identities = 243/346 (70%), Positives = 289/346 (83%) sp O54631 DDL_STRPN D-ALANINE--D-ALANINE LIGASE (D- ALANYLALANINE SYNTHETASE) (D-ALA-D-ALA LIGASE) gb AAC95435.1 (AF068901) D-Ala-D-Ala ligase [Streptococcus pneumoniae] Length = 347</p>
SeqID 1995	SA-627.1	Contig129 (35351-35947 m)	94	<p>Identities = 181/198 (91%), Positives = 189/198 (95%) sp P96053 RECR_STRTR RECOMBINATION PROTEIN RECR gb AAC44615.1 (U58210) RecM [Streptococcus thermophilus] Length = 198</p>
SeqID 1996	SA-628.1	Contig129 (35962-38004 m)	69	<p>Identities = 346/698 (49%), Positives = 482/698 (68%), Gaps = 17/698 (2%) gb AAC44614.1 (U58210) penicillin-binding protein 2b [Streptococcus thermophilus] Length = 704</p>

SeqID 1997	SA-629.1	Contig129 (38136-38828 m)	95	Identities = 219/230 (95%), Positives = 226/230 (98%) emb CAB51328.1 (AJ131985) phosphoglyceromutase [Streptococcus pneumoniae] Length = 230
SeqID 1998	SA-63.1	Contig131 (54046-54684 m)	16	Identities = 30/100 (30%), Positives = 43/100 (43%), Gaps = 21/100 (21%) gb AAG10259.1 AF264920_1 (AF264920) DS06238.4-like protein [Drosophila yakuba] Length = 213
SeqID 1999	SA-630.1	Contig129 (39005-39763 m)	77	Identities = 164/252 (65%), Positives = 202/252 (80%) sp P50918 TPIS_LACLA TRIOSEPHOSPHATE ISOMERASE (TIM) gb AAC43268.1 (U07640) triosephosphate isomerase [Lactococcus lactis] Length = 252
SeqID 2000	SA-632.1	Contig129 (40284-40760 p)	45	Identities = 48/145 (33%), Positives = 73/145 (50%), Gaps = 6/145 (4%) pir T39903 serine-rich protein - fission yeast (Schizosaccharomyces pombe) emb CAA22127.1 (AL033534) hypothetical serine-rich secreted protein [Schizosaccharomyces pombe] Length = 534
SeqID 2001	SA-633.1	Contig129 (39944-41140 m)	94	Identities = 384/398 (96%), Positives = 396/398 (99%) sp P33170 EFTU_STORR ELONGATION FACTOR TU (EF-TU) pir F60663 translation elongation factor EF-Tu - Streptococcus oralis Length = 398
SeqID 2002	SA-635.1	Contig129 (41492-42709 m)	53	Identities = 149/423 (35%), Positives = 229/423 (53%), Gaps = 41/423 (9%) sp Q47866 FTSW_ENTHR PROBABLE CELL DIVISION PROTEIN FTSW gb AAB39929.1 (U58049) putative cell division protein ftsW [Enterococcus hirae] Length = 397
SeqID 2003	SA-636.1	Contig129 (42866-45661 m)	59	Identities = 362/945 (38%), Positives = 559/945 (58%), Gaps = 65/945 (6%) sp O32483 CAPP_RHOA PHOSPHOENOLPYRUVATE CARBOXYLASE (PEPCASE) (PEPC) dbj BAA21835.1 (D89668) phosphoenolpyruvate carboxylase [Rhodospseudomonas palustris] Length = 936
SeqID 2004	SA-637.1	Contig129 (45870-47669 p)	80	Identities = 443/519 (85%), Positives = 489/519 (93%) dbj BAB16034.1 (AB030810) Bacillus licheniformis Pz-peptidase homologue [Streptococcus pyogenes] Length = 519

SeqID 2005	SA-638.1	Contig129 (47728-48117 m)	No Hits found	
SeqID 2006	SA-639.1	Contig129 (48101-48571 m)	No Hits found	
SeqID 2007	SA-64.1	Contig131 (53290-53961 m)	61	Identities = 102/221 (46%), Positives = 145/221 (65%) emb CAC07978.1 (AJ278983) CopR protein [Ralstonia metallidurans] Length = 228
SeqID 2008	SA-640.1	Contig129 (48875-49630 p)	32	Identities = 64/210 (30%), Positives = 95/210 (44%), Gaps = 15/210 (7%) emb CAA72266.1 (Y11477) endolysin [Bacteriophage Bastille] Length = 364
SeqID 2009	SA-641.1	Contig129 (49663-50280 m)	36	Identities = 38/156 (24%), Positives = 83/156 (52%), Gaps = 13/156 (8%) gb AAG20117.1 (AE005090) NADH dehydrogenase/oxidoreductase-like protein; NoIA [Halobacterium sp. NRC-1] Length = 303
SeqID 2010	SA-642.1	Contig129 (50575-51048 p)	59	Identities = 61/164 (37%), Positives = 96/164 (58%), Gaps = 13/164 (7%) sp P36922 EBSC_ENTFA EBSC PROTEIN pir C49939 ebsC protein - Enterococcus faecalis gb AAC36853.1 (L23802) regulatory protein [Enterococcus faecalis] Length = 164
SeqID 2011	SA-643.1	Contig129 (51058-51711 p)	42	Identities = 65/191 (34%), Positives = 93/191 (48%), Gaps = 13/191 (6%) pir G72260 phosphoglycerate mutase - Thermotoga maritima (strain MSB8) gb AAD36444.1 AE001791_6 (AE001791) phosphoglycerate mutase [Thermotoga maritima] Length = 201
SeqID 2012	SA-644.1	Contig129 (51747-52649 m)	34	Identities = 69/232 (29%), Positives = 108/232 (45%), Gaps = 9/232 (3%) pir E69814 conserved hypothetical protein yfnB - Bacillus subtilis dbj BAA20111.1 (D86418) YfnB [Bacillus subtilis] emb CAB12552.1 (Z99107) similar to hypothetical proteins [Bacillus subtilis] emb CAB12562.1 (Z99108) similar to hypothetical proteins [Bacillus subtilis] Length = 235

SeqID 2013	SA-646.1	Contig129 (52812-54314 p)	70	<p>Identities = 318/491 (64%), Positives = 391/491 (78%), Gaps = --1/491 (0%) sp P37477 SYK_BACSU LYSYL-TRNA SYNTHETASE (LYSINE--TRNA LIGASE) (LYSRS) pir S66111 lysine--tRNA ligase (EC 6.1.1.6) lysS - Bacillus subtilis dbj BAA05316.1 (D26185) lysyl-tRNA thynthetase [Bacillus subtilis] emb CAB11858.1 (Z99104) lysyl-tRNA synthetase [Bacillus subtilis] Length = 499</p>
SeqID 2014	SA-647.1	Contig129 (54389-54859 m)	69	<p>Identities = 103/151 (68%), Positives = 120/151 (79%) sp P11998 RISB_BACSU 6.7-DIMETHYL-8-RIBITYLLUMAZINE SYNTHASE (DMRL SYNTHASE) (LUMAZINE SYNTHASE) (RIBOFLAVIN SYNTHASE BETA CHAIN) pir A26708 riboflavin synthase (EC 2.5.1.9) complex beta chain ribH [validated] - Bacillus subtilis pdb 1RVV 1 Chain 1, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV 2 Chain 2, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV 3 Chain 3, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV 4 Chain 4, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV A Chain A, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV B Chain B, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV C Chain C, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV D Chain D, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV E Chain E, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS pdb 1RVV F Chain F, SynthaseRIBOFLAVIN SYNTHASE COMPLEX OF BACILLUS SUBTILIS</p>

SeqID 2015	SA-648.1	Contig129 (54874-56067 m)	72	<p>Identities = 230/395 (58%), Positives = 307/395 (77%) sp P50855 GCH2_ACTPL_RIBOFLAVIN_BIOSYNTHESIS PROTEIN_RIBA [INCLUDES: GTP CYCLOHYDROLASE II ; 3,4-DIHYDROXY-2-BUTANONE 4-PHOSPHATE SYNTHASE (DHBP SYNTHASE)] pir T50548 GTP cyclohydrolase II (EC 3.5.4.25) / 3, 4-dihydroxy-2-butanone 4-phosphate synthase (EC 5.4.99.-) [validated] - Actinobacillus pleuropneumoniae gb AAA86524.1 (U27202) GTP cyclohydrolase II/ 3,4-dihydroxy-2-butanone-4-phosphate synthase [Actinobacillus pleuropneumoniae] Length = 401</p>
SeqID 2016	SA-649.1	Contig129 (56085-56735 m)	67	<p>Identities = 98/216 (45%), Positives = 147/216 (67%), Gaps = 2/216 (0%) dbj BAB05274.1 (AP001512) riboflavin synthase alpha subunit [Bacillus halodurans] Length = 215</p>
SeqID 2017	SA-65.1	Contig131 (51942-53309 m)	49	<p>Identities = 114/466 (24%), Positives = 227/466 (48%), Gaps = 42/466 (9%) dbj BAB04092.1 (AP001508) two-component sensor histidine kinase [Bacillus halodurans] Length = 459</p>
SeqID 2018	SA-651.1	Contig129 (56716-57825 m)	66	<p>Identities = 180/353 (50%), Positives = 257/353 (71%) sp P50853 RIBD_ACTPL_RIBOFLAVIN_BIOSYNTHESIS PROTEIN_RIBD [INCLUDES: DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (HTP REDUCTASE)] Length = 376</p>
SeqID 2019	SA-653.1	Contig129 (58686-59894 p)	61	<p>Identities = 192/436 (44%), Positives = 274/436 (62%), Gaps = 14/436 (3%) pir C82733 manganese transport protein XF1015 [imported] - Xylella fastidiosa (strain 9a5c) gb AAAF83825.1 AE003939_4 (AE003939) manganese transport protein [Xylella fastidiosa] Length = 472</p>

SeqID 2020	SA-654.2	Contig129 (59996-60805 m)	40	Identities = 63/215 (29%), Positives = 108/215 (49%), Gaps = 13/215 (6%) pir H75355 hypothetical protein - Deinococcus radiodurans (strain R1) gb AAF11325.1 AE002018_7 (AE002018) hypothetical protein [Deinococcus radiodurans] Length = 250
SeqID 2021	SA-655.2	Contig129 (60917-61129 m)	53	Identities = 28/64 (43%), Positives = 41/64 (63%) sp O83371 Y352_TREPA HYPOTHETICAL PROTEIN TP0352 pir F71333 hypothetical protein TP0352 - syphilis spirochete gb AAC65352.1 (AE001215) T. pallidum predicted coding region TP0352 [Treponema pallidum] Length = 85
SeqID 2022	SA-656.2	Contig129 (61256-62542 m)	62	Identities = 201/407 (49%), Positives = 277/407 (67%), Gaps = 2/407 (0%) pir H69979 proteinase homolog yrrO - Bacillus subtilis emb CAB14676.1 (Z99117) similar to protease [Bacillus subtilis] Length = 422
SeqID 2023	SA-657.1	Contig129 (62672-63598 m)	50	Identities = 92/304 (30%), Positives = 161/304 (52%), Gaps = 5/304 (1%) pir G69979 proteinase homolog yrrN - Bacillus subtilis emb CAB14677.1 (Z99117) similar to protease [Bacillus subtilis] Length = 309
SeqID 2024	SA-658.2	Contig133 (43969-46152 m)	46	Identities = 294/409 (71%), Positives = 342/409 (82%), Gaps = 7/409 (1%) gb AAD00281.1 (U78600) putative ptsG protein [Streptococcus mutans] Length = 409
SeqID 2025	SA-659.1	Contig133 (43486-43767 p)	No Hits found	
SeqID 2026	SA-66.1	Contig131 (51118-51453 m)	48	Identities = 45/76 (59%), Positives = 54/76 (70%) pir T44087 hypothetical protein [imported] - Staphylococcus aureus (fragment) Length = 151
SeqID 2027	SA-660.1	Contig133 (43100-43915 m)	No Hits found	
SeqID 2028	SA-661.1	Contig133 (42071-42823 m)	71	Identities = 135/242 (55%), Positives = 183/242 (74%) emb CAC10170.1 (AJ278301) response regulator [Streptococcus pneumoniae] Length = 245
SeqID 2029	SA-662.1	Contig133 (41703-42002 m)	65	Identities = 44/88 (50%), Positives = 65/88 (73%) emb CAC10175.1 (AJ278302) histidine kinase [Streptococcus pneumoniae] Length = 446

SeqID 2030	SA-663.1	Contig133 (40732-41580 m)	55	Identities = 105/291 (36%), Positives = 161/291 (55%), Gaps = 18/291 (6%) emb[CAB88481.1] (AL353816) putative ABC transport system ATP-binding protein [Streptomyces coelicolor A3(2)] Length = 289
SeqID 2031	SA-664.1	Contig133 (39572-40726 m)	No Hits found	
SeqID 2032	SA-665.2	Contig133 (39693-39929 p)	No Hits found	
SeqID 2033	SA-666.1	Contig133 (38309-39376 m)	36	Identities = 78/262 (29%), Positives = 129/262 (48%), Gaps = 29/262 (11%) gb AAF71283.1 AF253562_7 (AF253562) racemase [Enterococcus faecalis] Length = 711
SeqID 2034	SA-667.1	Contig133 (37637-38014 m)	No Hits found	
SeqID 2035	SA-668.1	Contig133 (37144-37584 p)	No Hits found	
SeqID 2036	SA-669.1	Contig133 (37128-37619 m)	55	Identities = 49/164 (29%), Positives = 96/164 (57%), Gaps = 8/164 (4%) sp P37081 PTRB_KLEPN PTS SYSTEM, SORBOSE-SPECIFIC IIB COMPONENT (EIB-SOR) (SORBOSE-PERMEASE IIB COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, B COMPONENT) (EIII-B-SOR) pir S50188 phosphotransferase system enzyme II component B, sorbose-specific - Klebsiella pneumoniae emb CAA46858.1 (X66059) EIII-B Sor PTS [Klebsiella pneumoniae] prf 2022173D sorB gene [Klebsiella pneumoniae] Length = 164
SeqID 2037	SA-67.1	Contig131 (50330-51832 p)	74	Identities = 303/506 (59%), Positives = 389/506 (75%), Gaps = 6/506 (1%) sp P44023 YFCC_HAEIN HYPOTHETICAL PROTEIN HI0594 pir E64010 hypothetical protein HI0594 - Haemophilus influenzae (strain Rd KW20) gb AAC22251.1 (U32741) conserved hypothetical transmembrane protein [Haemophilus influenzae Rd] Length = 509
SeqID 2038	SA-670.1	Contig133 (36303-37112 m)	48	Identities = 70/251 (27%), Positives = 132/251 (51%), Gaps = 6/251 (2%) gb AAC44680.1 (U65015) PTS permease for mannose subunit IIPMan [Vibrio furnissii] Length = 258

SeqID 2039	SA-671.1	Contig133 (35479-36306 m)	54	<p>Identities = 94/280 (33%), Positives = 156/280 (55%), Gaps = 13/280 (4%) sp P08188 PTND_ECOLI PTS SYSTEM, MANNOSE-SPECIFIC IID COMPONENT (EIID-MAN) (MANNOSE-PERMEASE IID COMPONENT) (PHOSPHOTRANSFERASE ---- ENZYME II, D COMPONENT) (EII-M-MAN) pir WQECMM phosphotransferase system enzyme II (EC 2.7.1.69), mannose-specific, factor IID - Escherichia coli gb AAA24445.1 (J02699) mannose permease subunit II-M-Man [Escherichia coli] dbj BAA15631.1 (D90826) PTS system, Mannose-specific IID component (EIID-MAN) (Mannose- permease IID component) (Phosphotransferase enzyme II, D component) (EII-M-MAN). [Escherichia coli] gb AAC74889.1 (AE000276) PTS enzyme IID, mannose-specific [Escherichia coli K12] Length = 286</p>
SeqID 2040	SA-672.1	Contig133 (33692-35341 m)	24	<p>Identities = 84/258 (32%), Positives = 138/258 (52%), Gaps = 23/258 (8%) dbj BAB05628.1 (AP001513) two-component sensor histidine kinase [Bacillus halodurans] Length = 597</p>
SeqID 2041	SA-673.1	Contig133 (32915-33688 m)	56	<p>Identities = 86/252 (34%), Positives = 147/252 (58%), Gaps = 17/252 (6%) emb CAB54580.1 (AJ006398) response regulator [Streptococcus pneumoniae] gb AAF31452.1 AF221126_1 (AF221126) putative response regulator [Streptococcus pneumoniae] Length = 245</p>
SeqID 2042	SA-674.1	Contig133 (31866-32906 m)	53	<p>Identities = 114/331 (34%), Positives = 184/331 (55%), Gaps = 3/331 (0%) gb AAD18094.2 (U75349) periplasmic-iron-binding protein BtiA [Brachyspira hyodysenteriae] Length = 336</p>
SeqID 2043	SA-675.1	Contig133 (31147-31644 p)	46	<p>Identities = 55/150 (36%), Positives = 84/150 (55%), Gaps = 2/150 (1%) pir F75336 conserved hypothetical protein - Deinococcus radiodurans (strain R1) gb AAF11472.1 AE002031_1 (AE002031) conserved hypothetical protein [Deinococcus radiodurans] Length = 190</p>

SeqID 2044	SA-676.1	Contig133 (30110-31147 p)	70	Identities = 212/345 (61%), Positives = 268/345 (77%), Gaps = 1/345 (0%) pir [C69830 glucanase homolog ynfE - Bacillus subtilis emb CAA74526.1 (Y14083) hypothetical protein [Bacillus subtilis] emb CAB12860.1 (Z99109) similar to glucanase [Bacillus subtilis] Length = 346
SeqID 2045	SA-677.1	Contig133 (29644-30099 p)	80	Identities = 103/142 (72%), Positives = 123/142 (86%) emb CAB70606.1 (Y18363) ribonucleotide reductase-like (Nrd-like) protein [Streptococcus dysgalactiae subsp. equisimilis] Length = 142
SeqID 2046	SA-678.1	Contig133 (27085-29487 p)	61	Identities = 398/694 (57%), Positives = 493/694 (70%), Gaps = 30/694 (4%) emb CAB70615.1 (AJ133440) cyclo-nucleotide phosphodiesterase, putative [Streptococcus dysgalactiae subsp. equisimilis] Length = 683
SeqID 2047	SA-68.1	Contig131 (49259-50185 p)	65	Identities = 162/313 (51%), Positives = 207/313 (65%), Gaps = 7/313 (2%) sp Q46171 ARCC_CLOPE CARBAMATE KINASE emb CAA66367.1 (X97768) carbamate kinase [Clostridium perfringens] Length = 314
SeqID 2048	SA-680.1	Contig133 (24609-26825 m)	93	Identities = 647/739 (87%), Positives = 696/739 (93%), Gaps = 1/739 (0%) sp Q54089 RELA_STREQ PUTATIVE GTP PYROPHOSPHOKINASE (ATP:GTP 3 - PYROPHOSPHOTRANSFERASE) (PPGPP SYNTHETASE 1) ((P)PPGPP SYNTHETASE) (STRINGENT RESPONSE-LIKE PROTEIN) pir S39975 stringent response-like protein - Streptococcus equisimilis emb CAA51353.1 (X72832) stringent response-like protein [Streptococcus equisimilis] prf 2009358E stringent response-like protein [Streptococcus equisimilis] Length = 739
SeqID 2049	SA-681.1	Contig133 (24156-24599 m)	83	Identities = 104/145 (71%), Positives = 126/145 (86%) pir S39974 hypothetical protein - Streptococcus equisimilis emb CAA51352.1 (X72832) ORF1 [Streptococcus equisimilis] prf 2009358D rel upstream ORF [Streptococcus equisimilis] Length = 147

SeqID 2050	SA-682.1	Contig133 (22978-23901 m)	70	Identities = 180/304 (59%), Positives = 227/304 (74%), Gaps = 4/304 (1%) pir T46757 lipoprotein-lmb [validated] - Streptococcus agalactiae gb AAD13796.1 (AF062533) Lmb [Streptococcus agalactiae] Length = 306
SeqID 2051	SA-686.1	Contig133 (21506-22951 m)	55	Identities = 213/463 (46%), Positives = 277/463 (59%), Gaps = 41/463 (8%) pir T46758 hypothetical 92.4K protein - Streptococcus agalactiae gb AAD13797.1 (AF062533) unknown [Streptococcus agalactiae] Length = 822
SeqID 2052	SA-687.1	Contig133 (20370-21608 m)	24	Identities = 66/253 (26%), Positives = 101/253 (39%), Gaps = 64/253 (25%) pir T46758 hypothetical 92.4K protein - Streptococcus agalactiae gb AAD13797.1 (AF062533) unknown [Streptococcus agalactiae] Length = 822
SeqID 2053	SA-688.2	Contig133 (19312-20088 p)	68	Identities = 139/255 (54%), Positives = 189/255 (73%), Gaps = 6/255 (2%) gb AAA25174.1 (M35375) lactose repressor (lacR; alt.) [Lactococcus lactis] gb AAA25176.1 (M60447) repressor protein [Lactococcus lactis] gb AAA25186.1 (M60673) lacR [Lactococcus lactis] Length = 261
SeqID 2054	SA-689.2	Contig139 (185493-186197 p)	59	Identities = 99/228 (43%), Positives = 149/228 (64%), Gaps = 3/228 (1%) pir 40084 gtcR protein - Bacillus brevis emb CAA55264.1 (X78502) gtcR [Brevibacillus brevis] Length = 242
SeqID 2055	SA-69.1	Contig131 (48249-49247 p)	87	Identities = 264/332 (79%), Positives = 292/332 (87%) emb CAB75986.1 (AJ272085) ornithine carbamoyltransferase [Staphylococcus aureus] emb CAB75987.1 (AJ272086) ornithine carbamoyltransferase Otc6850 [Staphylococcus aureus] Length = 333
SeqID 2056	SA-690.1	Contig139 (183328-185271 p)	77	Identities = 412/638 (64%), Positives = 506/638 (78%), Gaps = 7/638 (1%) dbj BAB06860.1 (AP001517) threonyl-tRNA synthetase 1 [Bacillus halodurans] Length = 645
SeqID 2057	SA-691.1	Contig139 (181537-182871 p)	52	Identities = 189/290 (65%), Positives = 234/290 (80%) emb CAA72250.1 (Y11463) ORF5 [Streptococcus pneumoniae] Length = 290

SeqID 2058	SA-692.1	Contig139 (180537-181535 p)	53	Identities = 120/350 (34%), Positives = 180/350 (51%), Gaps = 29/350 (8%) gb AAG19110.1 (AE005009) Vng0600c [Halobacterium sp. NRC-1] Length = 361
SeqID 2059	SA-693.1	Contig139 (179026-180492 p)	76	Identities = 305/483 (63%), Positives = 378/483 (78%) gb AAC35010.1 (AF055987) intracellular α -amylase [Streptococcus mutans] Length = 486
SeqID 2060	SA-694.2	Contig139 (177890-178894 p)	94	Identities = 304/332 (91%), Positives = 320/332 (95%) dbj BAA88121.3 (AB028599) catabolite control protein A [Streptococcus bovis] Length = 334
SeqID 2061	SA-696.2	Contig139 (176595-177680 m)	81	Identities = 257/359 (71%), Positives = 304/359 (84%) gb AAC46293.1 (AF014460) PepQ [Streptococcus mutans] Length = 359
SeqID 2062	SA-697.1	Contig139 (174747-176537 p)	42	Identities = 148/586 (25%), Positives = 252/586 (42%), Gaps = 77/586 (13%) emb CAB66297.1 (AL136519) beta-N-acetylglucosaminidase. [Streptomyces coelicolor A3(2)] Length = 615
SeqID 2063	SA-698.1	Contig139 (173973-174731 p)	No Hits found	
SeqID 2064	SA-699.1	Contig139 (172927-173766 p)	76	Identities = 197/271 (72%), Positives = 228/271 (83%) sp P44481 Y048_HAEIN PUTATIVE OXIDOREDUCTASE HI0048 pir B64045 D-arabinitol 2-dehydrogenase homolog - Haemophilus influenzae (strain Rd KW20) gb AAC21726.1 (U32690) oxidoreductase [Haemophilus influenzae Rd] Length = 285
SeqID 2065	SA-7.1	Contig137 (38846-39100 m)	No Hits found	
SeqID 2066	SA-70.1	Contig131 (46851-48146 m)	44	Identities = 118/427 (27%), Positives = 199/427 (45%), Gaps = 25/427 (5%) dbj BAB06435.1 (AP001516) two-component sensor histidine kinase [Bacillus halodurans] Length = 437
SeqID 2067	SA-700.1	Contig139 (171757-172803 p)	73	Identities = 203/343 (59%), Positives = 261/343 (75%) dbj BAB04425.1 (AP001509) D-mannanase dehydratase [Bacillus halodurans] Length = 345

SeqID 2068	SA-701.1	Contig139 (170339-171739 p)	62	Identities = 215/465 (46%), Positives = 295/465 (63%), Gaps = 7/465 (1%) dbj BAB04424.1 (AP001509) uronate isomerase [Bacillus halodurans] Length = 472
SeqID 2069	SA-702.1	Contig139 (169705-170322 p)	51	Identities = 92/199 (46%), Positives = 124/199 (62%), Gaps = 6/199 (3%) pir F72422 KHG-KDPG bifunctional aldolase TM0066 [similarity] - Thermotoga maritima (strain MSB8) gb AAD35160.1 AE001693_6 (AE001693) 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase [Thermotoga maritima] Length = 205
SeqID 2070	SA-704.1	Contig139 (168917-169588 p)	50	Identities = 67/225 (29%), Positives = 119/225 (52%), Gaps = 17/225 (7%) sp P42239 YCBG_BACSU HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN GLTP-CWLJ INTERGENIC REGION (ORF6) pir B69753 transcription regulator GntR family homolog ycbG - Bacillus subtilis dbj BAA06471.1 (D30808) homologue of protein A which gene locates upstream pyruvate dehydrogenase gene cluster [Bacillus subtilis] emb CAB12044.1 (Z99105) similar to transcriptional regulator (GntR family) [Bacillus subtilis] Length = 233

SeqID 2071	SA-705.1	Contig139 (167089-168888 p)	57	Identities = 255/599 (42%), Positives = 356/599 (58%), Gaps = 25/599 (4%) gb AAF65327.1 (AF234293) GUSA, hexaHis tagged [Binary vector pCAMBIA-1201] gb AAF65330.1 (AF234294) GUSA, hexaHis tagged [Binary vector pCAMBIA-1281Z] gb AAF65334.1 (AF234295) GUSA, hexaHis tagged [Binary vector pCAMBIA-1291Z] gb AAF65342.1 (AF234297) GUSA, hexaHis tagged [Binary vector pCAMBIA-1301] gb AAF65372.1 (AF234306) GUSA, hexaHis tagged [Binary vector pCAMBIA-1381Z] gb AAF65390.1 (AF234312) GUSA, hexaHis tagged [Binary vector pCAMBIA-1391Z] gb AAF65397.1 (AF234314) GUSA, hexaHis tagged [Binary vector pCAMBIA-2201] gb AAF65404.1 (AF234316) GUSA, hexaHis tagged [Binary vector pCAMBIA-2301] Length = 620
SeqID 2072	SA-706.1	Contig139 (166047-167072 p)	49	Identities = 115/342 (33%), Positives = 179/342 (51%), Gaps = 16/342 (4%) pir G72422 2-keto-3-deoxygluconate kinase - Thermotoga maritima (strain MSB8) gb AAD35161.1 AE001693_7 (AE001693) 2-keto-3-deoxygluconate kinase [Thermotoga maritima] Length = 339
SeqID 2073	SA-707.1	Contig139 (164430-165980 p)	33	Identities = 107/443 (24%), Positives = 180/443 (40%), Gaps = 37/443 (8%) sp P94488 YNAJ_BACSU HYPOTHETICAL SYMPORTER IN GLNA-XYNB INTERGENIC REGION pir A69888 H+-symporter homolog ynaJ - Bacillus subtilis gb AAB41090.1 (U66480) YnaJ [Bacillus subtilis] emb CAB13641.1 (Z99113) similar to H+-symporter [Bacillus subtilis] Length = 463
SeqID 2074	SA-708.2	Contig139 (163229-164224 p)	68	Identities = 165/329 (50%), Positives = 230/329 (69%), Gaps = 1/329 (0%) gb AAB17663.1 (U31175) D-specific D-2-hydroxyacid dehydrogenase [Staphylococcus aureus] Length = 330
SeqID 2075	SA-71.1	Contig131 (46021-46854 m)	54	Identities = 100/305 (32%), Positives = 152/305 (49%), Gaps = 31/305 (10%) dbj BAB06434.1 (AP001516) two-component response regulator [Bacillus halodurans] Length = 312

SeqID 2076	SA-710.2	Contig139 (162160-163218 p)	55	Identities = 127/332 (38%), Positives = 202/332 (60%), Gaps = 1/332 (0%) pir E71373 probable regulatory protein (pfoS/R) - syphilis spirochete gb AAC65034.1 (AE001189) regulatory protein (pfoS/R) [Treponema pallidum] Length = 350
SeqID 2077	SA-712.1	Contig139 (161083-161967 m)	50	Identities = 88/282 (31%), Positives = 152/282 (53%), Gaps = 5/282 (1%) sp P39592 YWBI_BACSU HYPOTHETICAL TRANSCRIPTIONAL REGULATOR YWBI pir S39679 transcription regulator homolog ywbl - Bacillus subtilis emb CAA51580.1 (X73124) ipa-24d [Bacillus subtilis] emb CAB15857.1 (Z99123) alternate gene name: ipa-24d~similar to transcriptional regulator (LysR family) [Bacillus subtilis] Length = 301
SeqID 2078	SA-713.1	Contig139 (160228-161046 p)	No Hits found	
SeqID 2079	SA-714.1	Contig139 (159291-160064 p)	70	Identities = 128/253 (50%), Positives = 187/253 (73%) sp P77307 YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN USHA-TESA INTERGENIC REGION Length = 259
SeqID 2080	SA-715.1	Contig139 (158638-159294 p)	63	Identities = 95/202 (47%), Positives = 142/202 (70%), Gaps = 2/202 (0%) sp P77279 YBBL_ECOLI HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YBBL pir A64780 probable ABC transport protein ybbl - Escherichia coli gb AAB40244.1 (U82664) hypothetical protein [Escherichia coli] gb AAC73592.1 (AE000155) putative ATP-binding component of a transport system [Escherichia coli K12] Length = 225
SeqID 2081	SA-716.1	Contig139 (157783-158421 m)	59	Identities = 91/211 (43%), Positives = 131/211 (61%), Gaps = 7/211 (3%) sp P09548 DEDA_ECOLI DEDA PROTEIN (DSG-1 PROTEIN) pir XMECAD dedA protein - Escherichia coli gb AAA23964.1 (M68935) dedA [Escherichia coli] gb AAC75377.1 (AE000320) orf, hypothetical protein [Escherichia coli K12] dbj BAA16174.1 (D90863) dedA protein [Escherichia coli] Length = 219

SeqID 2082	SA-717.2	Contig139 (156980-157705 p)	50	Identities = 103/200 (51%), Positives = 136/200 (67%), Gaps = 4/200 (2%) sp Q03158 NUCE_STRPN DNA-ENTRY NUCLEASE (COMPETENCE-SPECIFIC NUCLEASE) pir S10641 endA protein - Streptococcus pneumoniae emb CAA38134.1 (X54225) membrane nuclease [Streptococcus pneumoniae] Length = 274
SeqID 2083	SA-719.1	Contig125 (39398-40735 m)	40	Identities = 151/256 (58%), Positives = 186/256 (71%), Gaps = 15/256 (5%) pir S72385 hypothetical protein 9 - Enterococcus faecalis plasmid pAD1 emb CAA65668.1 (X96977) orf9 [Enterococcus faecalis] Length = 278
SeqID 2084	SA-72.1	Contig131 (45077-45799 m)	72	Identities = 139/236 (58%), Positives = 178/236 (74%) gb AAD45529.1 AF162655_1 (AF162655) choline transporter [Streptococcus pneumoniae] Length = 242
SeqID 2085	SA-720.1	Contig125 (38800-39384 m)	24	Identities = 36/142 (25%), Positives = 62/142 (43%), Gaps = 7/142 (4%) pir D70463 hypothetical protein aq_1894 - Aquifex aeolicus gb AAC07706.1 (AE000762) putative protein [Aquifex aeolicus] Length = 237
SeqID 2086	SA-721.1	Contig125 (37843-38661 m)	53	Identities = 89/267 (33%), Positives = 157/267 (58%), Gaps = 15/267 (5%) gb AAF27301.1 AF154674_2 (AF154674) ParA [Lactococcus lactis subsp. lactis] gb AAF27325.1 (AF178424) ParA [Lactococcus lactis] Length = 262
SeqID 2087	SA-722.1	Contig125 (37568-37846 m)	No Hits found	
SeqID 2088	SA-723.1	Contig125 (37172-37561 m)	40	Identities = 33/93 (35%), Positives = 53/93 (56%), Gaps = 2/93 (2%) ref NP_038304.1 ORF4 [Streptococcus thermophilus bacteriophage 7201] gb AAF26603.1 AF118440_1 (AF145054) ORF4 [Streptococcus thermophilus bacteriophage 7201] Length = 268
SeqID 2089	SA-725.1	Contig125 (36856-37167 m)	No Hits found	
SeqID 2090	SA-726.1	Contig125 (35394-36722 m)	No Hits found	
SeqID 2091	SA-727.1	Contig125 (34305-35003 m)	66	Identities = 106/222 (47%), Positives = 162/222 (72%), Gaps = 2/222 (0%) gb AAG28336.1 (U88582) SatD [Streptococcus mutans] Length = 222

SeqID 2092	SA-728.1	Contig125 (33542-34177 m)	26	Identities = 54/103 (52%), Positives = 70/103 (67%), Gaps = 2/103 (1%) gb AAG28337.1 (U88582) SatE [Streptococcus mutans] Length = 108
SeqID 2093	SA-729.1	Contig125 (31914-33509 p)	75	Identities = 346/521 (66%), Positives = 411/521 (78%), Gaps = 24/521 (4%) gb AAD17886.1 (AF100456) hyaluronate-associated protein precursor [Streptococcus equi] Length = 522
SeqID 2094	SA-73.1	Contig131 (44407-44805 p)	No Hits found	
SeqID 2095	SA-730.1	Contig125 (31012-31461 m)	83	Identities = 119/148 (80%), Positives = 136/148 (91%) gb AAC17173.1 (AF065141) unknown [Streptococcus mutans] Length = 356
SeqID 2096	SA-731.1	Contig125 (30392-31078 m)	89	Identities = 188/218 (86%), Positives = 205/218 (93%) gb AAC17173.1 (AF065141) unknown [Streptococcus mutans] Length = 356
SeqID 2097	SA-732.1	Contig125 (29462-30349 p)	52	Identities = 96/320 (30%), Positives = 172/320 (53%), Gaps = 16/320 (5%) sp P42422 YXDK_BACSU HYPOTHETICAL SENSOR-LIKE HISTIDINE KINASE IN IDH 3 REGION pir H70073 two-component sensor histidine kinase homolog yxdK - Bacillus subtilis dbj BAA03301.1 (D14399) hypothetical protein [Bacillus subtilis] emb CAB16001.1 (Z99124) similar to two-component sensor histidine kinase [YxdJ] [Bacillus subtilis] Length = 325
SeqID 2098	SA-733.1	Contig125 (28746-29414 p)	55	Identities = 95/226 (42%), Positives = 136/226 (60%), Gaps = 10/226 (4%) pir D70032 two-component response regulator [YvcQ] homolog yvcP - Bacillus subtilis emb CAB08062.1 (Z94043) hypothetical protein [Bacillus subtilis] emb CAB15477.1 (Z99121) similar to two-component response regulator [YvcQ] [Bacillus subtilis] Length = 237
SeqID 2099	SA-734.1	Contig125 (26819-28636 p)	45	Identities = 154/666 (23%), Positives = 300/666 (44%), Gaps = 40/666 (6%) gb AAF99695.1 AF267498_5 (AF267498) permease OrfY [Streptococcus mutans] Length = 640

SeqID 2100	SA-735.1	Contig125 (25927-26679 p)	68	Identities = 118/242 (48%), Positives = 175/242 (71%), Gaps = 1/242 (0%) gb AAAF99694.1 AF267498.4 (AF267498) ABC transporter OrfX [Streptococcus mutans] Length = 246
SeqID 2101	SA-736.1	Contig125 (25443-25901 p)	56	Identities = 59/146 (40%), Positives = 87/146 (59%), Gaps = 3/146 (2%) gb AAB08491.1 (U25181) nisin-resistance protein [Lactococcus lactis] Length = 318
SeqID 2102	SA-737.1	Contig125 (25003-25479 p)	52	Identities = 51/177 (28%), Positives = 104/177 (57%), Gaps = 5/177 (2%) gb AAB08491.1 (U25181) nisin-resistance protein [Lactococcus lactis] Length = 318
SeqID 2103	SA-738.1	Contig125 (24286-24774 m)	No Hits found	
SeqID 2104	SA-739.1	Contig125 (23406-24254 p)	62	Identities = 123/290 (42%), Positives = 178/290 (60%), Gaps = 18/290 (6%) emb CAB59827.1 (AJ012388) hypothetical protein [Lactococcus lactis] Length = 286
SeqID 2105	SA-74.1	Contig131 (43560-45074 m)	80	Identities = 326/506 (64%), Positives = 410/506 (80%), Gaps = 1/506 (0%) pir T44634 choline transporter [imported] - Streptococcus pneumoniae gb AAD45530.1 AF162656_1 (AF162656) choline transporter [Streptococcus pneumoniae] Length = 506
SeqID 2106	SA-740.1	Contig125 (22530-23273 p)	No Hits found	
SeqID 2107	SA-741.1	Contig125 (21118-22452 p)	79	Identities = 291/439 (66%), Positives = 353/439 (80%), Gaps = 10/439 (2%) sp P39815 GID_BACSU GID PROTEIN pir A69632 glucose-inhibited division protein gid - Bacillus subtilis emb CAA04423.1 (AJ000975) Gid protein [Bacillus subtilis] emb CAB13486.1 (Z99112) glucose-inhibited division protein [Bacillus subtilis] Length = 435
SeqID 2108	SA-742.1	Contig125 (20306-21004 p)	56	Identities = 83/229 (36%), Positives = 133/229 (57%), Gaps = 1/229 (0%) dbj BAB04138.1 (AP001508) transcriptional regulator (GntR family) [Bacillus halodurans] Length = 240

SeqID 2109	SA-743.2	Contig125 (18531-20093 m)	86	Identities = 418/511 (81%), Positives = 467/511 (90%), Gaps = 3/511 (0%) gb AAK05584.1 AE006379_1 (AE006379) GMP synthase (EC 6.3.5.2) [Lactococcus lactis subsp. lactis] Length = 513
SeqID 2110	SA-744.2	Contig119 (32996-33796 m)	57	Identities = 113/251 (45%), Positives = 160/251 (63%), Gaps = 3/251 (1%) gb AAK04846.1 AE006308_6 (AE006308) HYPOTHETICAL PROTEIN [Lactococcus lactis subsp. lactis] Length = 257
SeqID 2111	SA-745.1	Contig119 (32371-33006 m)	74	Identities = 126/211 (59%), Positives = 161/211 (75%) pir B69997 conserved hypothetical protein ytmQ - Bacillus subtilis gb AAC00285.1 (AF008220) YtmQ [Bacillus subtilis] emb CAB14968.1 (Z99119) similar to hypothetical proteins [Bacillus subtilis] Length = 213
SeqID 2112	SA-746.1	Contig119 (31405-31899 m)	60	Identities = 73/152 (48%), Positives = 99/152 (65%), Gaps = 3/152 (1%) dbj BAB06136.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 156
SeqID 2113	SA-747.1	Contig119 (30218-31369 m)	63	Identities = 164/370 (44%), Positives = 251/370 (67%), Gaps = 15/370 (4%) sp P32727 NUSA_BACSU N UTILIZATION SUBSTANCE PROTEIN A HOMOLOG emb CAB13533.1 (Z99112) nusa [Bacillus subtilis] Length = 371
SeqID 2114	SA-748.1	Contig119 (29900-30196 m)	65	Identities = 46/92 (50%), Positives = 67/92 (72%), Gaps = 1/92 (1%) sp P32728 YLXR_BACSU HYPOTHETICAL 10.4 KD PROTEIN IN NUSA-INFN INTERGENIC REGION (ORF3) pir D36905 conserved hypothetical protein ylxR - Bacillus subtilis emb CAA79232.1 (Z18631) ORF3 [Bacillus subtilis] emb CAB13534.1 (Z99112) alternate gene name: ymx8~similar to hypothetical proteins [Bacillus subtilis] Length = 91
SeqID 2115	SA-749.1	Contig119 (29605-29907 m)	67	Identities = 50/97 (51%), Positives = 69/97 (70%) sp P55768 YLXQ_ENTFC PROBABLE RIBOSOMAL PROTEIN IN INFN 5 REGION Length = 103
SeqID 2116	SA-75.1	Contig131 (40941-43481 p)	38	Identities = 190/596 (31%), Positives = 324/596 (53%), Gaps = 31/596 (5%) dbj BAA24464.1 (D85082) YfiX [Bacillus subtilis] Length = 610

SeqID 2117	SA-751.1	Contig119 (26802-29585 m)	98	<p>Identities = 926/927 (99%), Positives = 927/927 (99%)</p> <p>sp Q9ZF20 IF2_STRAG TRANSLATION INITIATION FACTOR IF-2 emb CAA05919.1 (AJ003164) initiation factor IF2 [Streptococcus agalactiae] emb CAC00489.1 (AJ251495) initiation factor 2 [Streptococcus agalactiae] emb CAC00491.1 (AJ251496) initiation factor 2 [Streptococcus agalactiae]</p> <p>Length = 927</p>
SeqID 2118	SA-752.1	Contig119 (26343-26711 m)	91	<p>Identities = 122/122 (100%), Positives = 122/122 (100%)</p> <p>emb CAA05920.1 (AJ003164) ribosome binding factor A [Streptococcus agalactiae] emb CAC00486.1 (AJ251493) ribosome binding factor A [Streptococcus agalactiae] emb CAC00488.1 (AJ251494) ribosome binding factor A [Streptococcus agalactiae] emb CAC00490.1 (AJ251495) ribosome binding factor A [Streptococcus agalactiae] emb CAC00492.1 (AJ251496) ribosome binding factor A [Streptococcus agalactiae] emb CAC00498.1 (AJ251499) ribosome binding factor A [Streptococcus agalactiae]</p> <p>Length = 122</p>
SeqID 2119	SA-753.1	Contig119 (25234-26259 p)	26	<p>Identities = 58/220 (26%), Positives = 90/220 (40%), Gaps = 8/220 (3%) sp Q01109 BAH_STRHY ACETYL-HYDROLASE gb AAA79277.1 (M64783) acetyl-hydrolase [Streptomyces hygroscopicus] Length = 299</p>
SeqID 2120	SA-754.1	Contig119 (24675-25091 m)	68	<p>Identities = 67/138 (48%), Positives = 99/138 (71%)</p> <p>gb AAG10085.1 AF296446_1 (AF296446) CopY [Streptococcus mutans] Length = 147</p>
SeqID 2121	SA-756.1	Contig119 (22428-24662 m)	76	<p>Identities = 440/740 (59%), Positives = 571/740 (76%), Gaps = 1/740 (0%) gb AAG10086.1 AF296446_2 (AF296446) CopA [Streptococcus mutans] Length = 742</p>
SeqID 2122	SA-757.1	Contig119 (22181-22387 m)	58	<p>Identities = 31/67 (46%), Positives = 43/67 (63%)</p> <p>gb AAG10087.1 AF296446_3 (AF296446) CopZ [Streptococcus mutans] Length = 67</p>

SeqID 2123	SA-758.1	Contig119 (21457-22071 m)	65	Identities = 107/192 (55%), Positives = 137/192 (70%) sp O32216 VVG_T-BAGSU-HYPOTHETICAL 21.7 KDA PROTEIN IN SSPG-SECG INTERGENIC REGION pir A70041 conserved hypothetical protein yvgT - Bacillus subtilis emb CAB15351.1 (Z99121) similar to hypothetical proteins [Bacillus subtilis] Length = 202
SeqID 2124	SA-759.1	Contig119 (20630-21460 m)	66	Identities = 138/238 (57%), Positives = 184/238 (76%) emb CAB94816.1 (AJ245582) hypothetical protein [Streptococcus thermophilus] Length = 240
SeqID 2125	SA-76.1	Contig131 (40208-40957 p)	44	Identities = 58/255 (22%), Positives = 115/255 (44%), Gaps = 26/255 (10%) ref NP_069699.1 carboxylesterase (est-1) [Archaeoglobus fulgidus] pir A69358 carboxylesterase (est-1) homolog - Archaeoglobus fulgidus gb AAB90371.1 (AE001044) carboxylesterase (est-1) [Archaeoglobus fulgidus] Length = 266
SeqID 2126	SA-760.1	Contig119 (17875-20517 m)	83	Identities = 655/879 (74%), Positives = 748/879 (84%), Gaps = 4/879 (0%) sp P13252 DPO1_STRPN DNA POLYMERASE I (POL I) pir A32949 DNA-directed DNA polymerase (EC 2.7.7.7) - Streptococcus pneumoniae gb AAA26954.1 (J04479) DNA polymerase I [Streptococcus pneumoniae] Length = 877
SeqID 2127	SA-761.1	Contig119 (17405-17845 m)	58	Identities = 72/134 (53%), Positives = 94/134 (69%), Gaps = 3/134 (2%) db BAB05860.1 (AP001514) unknown conserved protein [Bacillus halodurans] Length = 137
SeqID 2128	SA-762.1	Contig119 (16844-17323 m)	75	Identities = 113/156 (72%), Positives = 135/156 (86%), Gaps = 1/156 (0%) gb AAB18795.1 (U76538) Fur-like protein [Streptococcus pyogenes] Length = 155
SeqID 2129	SA-765.1	Contig119 (15126-16691 m)	11	Identities = 41/146 (28%), Positives = 63/146 (43%), Gaps = 13/146 (8%) emb CAA06650.1 (AJ005645) sdrc [Staphylococcus aureus] Length = 947

SeqID 2130	SA-766.1	Contig119 (14327-15013 m)	-78-	Identities = 143/228 (62%), Positives = 183/228 (79%), Gaps = 1/228 (0%) emb CAB54578.1 (AJ006397)-response-regulator [Streptococcus pneumoniae] Length = 232
SeqID 2131	SA-767.1	Contig119 (13288-14325 m)	71	Identities = 190/343 (55%), Positives = 249/343 (72%) emb CAB54579.1 (AJ006397) histidine kinase [Streptococcus pneumoniae] Length = 350
SeqID 2132	SA-768.2	Contig119 (12486-13274 p)	23	Identities = 35/119 (29%), Positives = 62/119 (51%), Gaps = 15/119 (12%) pir S32215 hypothetical protein 1 - Bacillus megaterium emb CAA79984.1 (Z21972) ORF1 [Bacillus megaterium] Length = 226
SeqID 2133	SA-769.2	Contig119 (11205-12347 m)	80	Identities = 269/377 (71%), Positives = 320/377 (84%) sp O32053 TGT_BACSU QUEUINE TRNA-RIBOSYLTRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME) pir B69722 queuine tRNA-ribosyltransferase (EC 2.4.2.29) - Bacillus subtilis emb CAB14731.1 (Z99118) tRNA-guanine transglycosylase [Bacillus subtilis] emb CAB75333.1 (Y15896) tRNA-guanine transglycosylase [Bacillus subtilis] Length = 381
SeqID 2134	SA-77.1	Contig131 (39724-40080 p)	68	Identities = 70/119 (58%), Positives = 89/119 (73%), Gaps = 1/119 (0%) sp Q02009 YTRP_LACLA HYPOTHETICAL 13.3 KD PROTEIN IN TRPE 5 REGION pir S35123 hypothetical protein (trpE 5 region) - Lactococcus lactis subsp. lactis gb AAA25222.1 (M87483) ORF 1 [Lactococcus lactis] Length = 119
SeqID 2135	SA-770.1	Contig119 (10787-11098 m)	59	Identities = 45/98 (45%), Positives = 65/98 (65%), Gaps = 3/98 (3%) pir T38499 hypothetical protein SPAC29B12.12 - fission yeast (Schizosaccharomyces pombe) emb CAB16256.1 (Z99164) hypothetical protein [Schizosaccharomyces pombe] Length = 113

SeqID 2136	SA-772.1	Contig119 (10241-10780 m)	53	Identities = 62/173 (35%), Positives = 97/173 (55%), Gaps = 1/173 (0%) pir [B69832 biotin biosynthesis homolog yhfU - Bacillus subtilis emb CAA74544.1 (Y14084) hypothetical protein [Bacillus subtilis] emb CAB12877.1 (Z99109) similar to biotin biosynthesis [Bacillus subtilis] Length = 186
SeqID 2137	SA-773.1	Contig119 (9326-10102 m)	48	Identities = 89/218 (40%), Positives = 129/218 (58%), Gaps = 4/218 (1%) pir [F69830 conserved hypothetical protein yhfI - Bacillus subtilis emb CAA74530.1 (Y14083) hypothetical protein [Bacillus subtilis] emb CAB12864.1 (Z99109) similar to hypothetical proteins [Bacillus subtilis] Length = 244
SeqID 2138	SA-774.2	Contig119 (8820-9326 m)	62	Identities = 71/154 (46%), Positives = 110/154 (71%) sp P21335 YAAJ_BACSU HYPOTHETICAL 17.8 KD PROTEIN IN SERS-DNAH INTERGENIC REGION pir [S11690 conserved hypothetical protein yaaJ - Bacillus subtilis emb CAA36389.1 (X52144) ORF17 (AA 1-161) [Bacillus subtilis] dbj BAA05254.1 (D26185) unknown [Bacillus subtilis] emb CAB11794.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] prf 1617102A 17kD protein [Bacillus subtilis] Length = 161
SeqID 2139	SA-776.1	Contig110 (3335-3646 p)	No Hits found	
SeqID 2140	SA-777.1	Contig110 (2062-2373 m)	27	Identities = 25/35 (71%), Positives = 29/35 (82%) pir [G81516 hypothetical protein CP0988 [imported] - Chlamydomonas pneumoniae (strain AR39) Length = 52
SeqID 2141	SA-778.1	Contig131 (3258-3365 p)	No Hits found	
SeqID 2142	SA-779.1	Contig110 (1371-1676 m)	No Hits found	

SeqID 2143	SA-78.1	Contig131 (38025-39656 p)	48	<p>Identities = 159/568 (27%), Positives = 265/568 (45%), Gaps = 82/568 (14%) sp P32399 YHGE_BACSU HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB) pir H69832 phage infection protein homolog yhgE - Bacillus subtilis emb CAA74522.1 (Y14083) hypothetical protein [Bacillus subtilis] emb CAB12856.1 (Z99109) alternate gene name: yixE~similar to phage infection protein [Bacillus subtilis] Length = 775</p>
SeqID 2144	SA-780.1	Contig110 (499-846 m)	No Hits found	
SeqID 2145	SA-782.2	Contig139 (106281-108005 p)	46	<p>Identities = 122/553 (22%), Positives = 265/553 (47%), Gaps = 12/553 (2%) dbj BAB06924.1 (AP001518) BH3205~unknown conserved protein [Bacillus halodurans] Length = 561</p>
SeqID 2146	SA-783.1	Contig139 (108099-108740 p)	64	<p>Identities = 94/212 (44%), Positives = 137/212 (64%) pir E70860 probable serB2 protein - Mycobacterium tuberculosis (strain H37RV) emb CAA16127.1 (AL021287) serB2 [Mycobacterium tuberculosis] Length = 409</p>
SeqID 2147	SA-784.1	Contig139 (108761-109246 m)	24	<p>Identities = 24/77 (31%), Positives = 40/77 (51%) gb AAG20030.1 (AE005083) isopentenyl pyrophosphate isomerase; Idi [Halobacterium sp. NRC-1] Length = 213</p>
SeqID 2148	SA-785.1	Contig139 (109259-109714 m)	36	<p>Identities = 32/116 (27%), Positives = 57/116 (48%), Gaps = 4/116 (3%) pir B70008 hypothetical protein yuel - Bacillus subtilis emb CAB15165.1 (Z99120) yuel [Bacillus subtilis] Length = 132</p>
SeqID 2149	SA-786.1	Contig139 (109912-111219 p)	92	<p>Identities = 396/435 (91%), Positives = 414/435 (95%), Gaps = 1/435 (0%) dbj BAA81815.1 (AB029313) enclose [Streptococcus intermedius] Length = 434</p>
SeqID 2150	SA-787.1	Contig139 (111327-112391 m)	30	<p>Identities = 62/236 (26%), Positives = 107/236 (45%), Gaps = 38/236 (16%) gb AAF20148.1 (AF208390_1) (AF208390) actinin-like protein [Entamoeba histolytica] Length = 537</p>

SeqID 2151	SA-788.1	Contig139 (112620-113903 p)	79	<p>Identities = 287/426 (67%), Positives = 346/426 (80%) sp Q9S400 AROA_STRPN_3-PHOSPHOSHIKIMATE_1-CARBOXYVINYLTRANSFERASE (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS) gb AAD45819.1 AF169483_1 (AF169483) 5-enolpyruvylshikimate-3-phosphate synthase [Streptococcus pneumoniae] Length = 427</p>
SeqID 2152	SA-789.1	Contig139 (113896-114408 p)	56	<p>Identities = 65/164 (39%), Positives = 98/164 (59%), Gaps = 8/164 (4%) sp P43906 AROK_LACLA_SHIKIMATE_KINASE (SK pir) S52581 shikimate kinase (EC 2.7.1.71) - Lactococcus lactis emb CAA55181.1 (X78413) shikimate kinase [Lactococcus lactis] Length = 162</p>
SeqID 2153	SA-79.1	Contig131 (37307-38101 p)	51	<p>Identities = 109/229 (47%), Positives = 165/229 (71%) gb AAC14610.1 (U95842) transmembrane protein Tmp7 [Lactococcus lactis] Length = 234</p>
SeqID 2154	SA-790.1	Contig139 (114432-115838 p)	43	<p>Identities = 143/409 (34%), Positives = 206/409 (49%), Gaps = 56/409 (13%) emb CAB76821.1 (AJ276231) PSR protein [Enterococcus faecalis] emb CAB76822.1 (AJ276232) PSR protein [Enterococcus faecalis] Length = 390</p>
SeqID 2155	SA-792.1	Contig139 (115939-117294 p)	63	<p>Identities = 198/452 (43%), Positives = 300/452 (65%) dbj BAB04406.1 (AP001509) RNA methyltransferase [Bacillus halodurans] Length = 458</p>
SeqID 2156	SA-793.1	Contig139 (117321-117623 p)	No Hits found	
SeqID 2157	SA-794.1	Contig139 (117741-118478 p)	58	<p>Identities = 112/245 (45%), Positives = 148/245 (59%), Gaps = 10/245 (4%) pir T09434 acid phosphatase (EC 3.1.3.2) [similarity] - Haemophilus influenzae (strains CCUG, Rd KW20) emb CAA68889.1 (Y07615) acid phosphatase [Haemophilus influenzae] Length = 235</p>

SeqID 2158	SA-795.1	Contig139 (118799-119317 p)	67	Identities = 93/173 (53%), Positives = 123/173 (70%), Gaps = 3/173 (1%) ref NP_010829.1 Ydr540cp [Saccharomyces cerevisiae] pir S62019 hypothetical protein YDR540c - yeast (Saccharomyces cerevisiae) gb AAB64982.1 (U43834) Ydr540cp [Saccharomyces cerevisiae] Length = 179
SeqID 2159	SA-796.1	Contig139 (119446-119634 m)	No Hits found	
SeqID 2160	SA-797.1	Contig139 (119663-120118 m)	32	Identities = 29/95 (30%), Positives = 49/95 (51%), Gaps = 2/95 (2%) gb AAG09969.1 AF248037_4 (AF248037) unknown [Streptococcus agalactiae] Length = 174
SeqID 2161	SA-798.1	Contig139 (120154-120486 p)	34	Identities = 31/63 (49%), Positives = 44/63 (69%), Gaps = 2/63 (3%) gb AAG01392.1 AF208158_1 (AF208158) alpha-like protein 2 [Streptococcus agalactiae] Length = 786
SeqID 2162	SA-799.1	Contig139 (120607-121428 m)	92	Identities = 285/296 (96%), Positives = 287/296 (96%) pir T09011 probable transposase TnpA - Streptococcus pyogenes (fragment) gb AAB92607.1 (AF026542) TnpA [Streptococcus pyogenes] Length = 364
SeqID 2163	SA-8.1	Contig137 (36802-39009 p)	28	Identities = 123/560 (21%), Positives = 212/560 (36%), Gaps = 74/560 (13%) ref NP_053169.1 pXO2-14 [Bacillus anthracis] gb AAF13619.1 AF188935_17 (AF188935) pXO2-14 [Bacillus anthracis] Length = 952
SeqID 2164	SA-80.1	Contig131 (36659-37198 m)	51	Identities = 47/181 (25%), Positives = 95/181 (51%), Gaps = 16/181 (8%) dbj BAB04438.1 (AP001509) transcriptional regulator (TetR/AcrR family) [Bacillus halodurans] Length = 188
SeqID 2165	SA-800.1	Contig139 (121476-121751 m)	61	Identities = 66/80 (82%), Positives = 68/80 (84%) pir T09011 probable transposase TnpA - Streptococcus pyogenes (fragment) gb AAB92607.1 (AF026542) TnpA [Streptococcus pyogenes] Length = 364

SeqID 2166	SA-801.1	Contig139 (121884-122021 p)	78	Identities = 36/49 (73%), Positives = 40/49 (81%) gb AAC98436.1 (L29324) unknown [Streptococcus pneumoniae] Length = 118
SeqID 2167	SA-803.1	Contig139 (122409-122756 m)	82	Identities = 93/98 (94%), Positives = 96/98 (97%) sp Q9ZB45 HLSO_STRPY_33 KDA CHAPERONIN (HEAT SHOCK PROTEIN 33 HOMOLOG) (HSP33) gb AAC97154.1 (U49397) unknown [Streptococcus pyogenes] Length = 277
SeqID 2168	SA-804.1	Contig139 (122950-123369 m)	85	Identities = 104/135 (77%), Positives = 119/135 (88%) gb AAC97150.1 (U49397) MsmR [Streptococcus pyogenes] Length = 209
SeqID 2169	SA-805.1	Contig139 (123370-124158 m)	23	Identities = 59/72 (81%), Positives = 66/72 (90%) gb AAC97150.1 (U49397) MsmR [Streptococcus pyogenes] Length = 209
SeqID 2170	SA-806.1	Contig139 (124541-126205 p)	37	Identities = 141/512 (27%), Positives = 222/512 (42%), Gaps = 94/512 (18%) pir S52348 hypothetical protein 2 - Lactobacillus leichmannii emb CAA57459.1 (X81869) orf2 [Lactobacillus leichmannii] Length = 507
SeqID 2171	SA-807.1	Contig139 (126330-127217 p)	20	Identities = 45/141 (31%), Positives = 63/141 (43%), Gaps = 20/141 (14%) dbj BAB04080.1 (AP001508) unknown [Bacillus halodurans] Length = 1661
SeqID 2172	SA-808.1	Contig139 (127219-128136 p)	46	Identities = 96/265 (36%), Positives = 150/265 (56%), Gaps = 10/265 (3%) gb AAC13546.1 (AF019629) putative fimbria- associated protein [Actinomyces naeslundii] Length = 365
SeqID 2173	SA-809.3	Contig139 (128192-128944 p)	36	Identities = 79/178 (44%), Positives = 112/178 (62%), Gaps = 7/178 (3%) gb AAC13546.1 (AF019629) putative fimbria- associated protein [Actinomyces naeslundii] Length = 365
SeqID 2174	SA-81.1	Contig131 (36277-36573 p)	66	Identities = 46/97 (47%), Positives = 69/97 (70%) ref NP_053020.1 hypothetical protein [Plasmid pNZ4000] gb AAD40361.1 (AF036485) hypothetical protein [Plasmid pNZ4000] gb AAF98302.1 (AF243383_3 (AF243383) unknown; Orf3 [Lactococcus lactis subsp. lactis] Length = 98

SeqID 2175	SA-810.3	Contig139 (128806-128955 p)	No Hits found	
SeqID 2176	SA-811.2	Contig123 (32704-32853 p)	85	Identities = 43/49 (87%), Positives = 46/49 (93%) sp O34102 RL33_LACLC 50S RIBOSOMAL PROTEIN L33 gb AAB66692.1 (U89998) 50S ribosomal protein subunit L33 [Lactococcus lactis subsp. cremoris] Length = 49
SeqID 2177	SA-812.1	Contig123 (32506-32688 p)	55	Identities = 46/53 (86%), Positives = 49/53 (91%) sp O34101 RL32_LACLC 50S RIBOSOMAL PROTEIN L32 gb AAB66691.1 (U89998) 50S ribosomal protein subunit L32 [Lactococcus lactis subsp. cremoris] Length = 58
SeqID 2178	SA-814.1	Contig123 (31006-32286 m)	84	Identities = 342/423 (80%), Positives = 377/423 (88%) sp P30053 SYH_STREQ HISTIDYL-TRNA SYNTHETASE (HISTIDINE--TRNA LIGASE) (HISRS) pir S30233 histidine-- tRNA ligase (EC 6.1.1.21) [validated] - Streptococcus equisimilis Length = 426
SeqID 2179	SA-817.1	Contig123 (29159-30913 m)	73	Identities = 339/586 (57%), Positives = 432/586 (72%), Gaps = 10/586 (1%) sp O32038 SYD_BACSU ASPARTYL-TRNA SYNTHETASE (ASPARTATE--TRNA LIGASE) (ASPRS) pir D69591 aspartate--tRNA ligase (EC 6.1.1.12) aspS - Bacillus subtilis emb CAB14714.1 (Z99118) aspartyl-tRNA synthetase [Bacillus subtilis] Length = 592
SeqID 2180	SA-819.1	Contig123 (28225-29175 m)	57	Identities = 105/275 (38%), Positives = 182/275 (66%), Gaps = 1/275 (0%) emb CAB01834.1 (Z79580) putative ORF [Bacillus subtilis] emb CAA70630.1 (Y09476) YitT [Bacillus subtilis] emb CAB12952.1 (Z99109) alternate gene name: yuxA~similar to hypothetical proteins [Bacillus subtilis] Length = 280

SeqID 2181	SA-82.1	Contig131 (35422-36033 p)	77	Identities = 138/201 (68%), Positives = 158/201 (77%), Gaps = 1/201 (0%) sp P21466 RS4_BACSU_30S_RIBOSOMAL_PROTEIN_S4 [BS4] pir A37146 ribosomal protein S4 - Bacillus subtilis gb AAA22717.1 (M59358) ribosomal protein S4 [Bacillus subtilis] gb AAC00397.1 (AF008220) ribosomal protein S4 [Bacillus subtilis] emb CAB14944.1 (Z99119) ribosomal protein S4 (BS4) [Bacillus subtilis] Length = 200
SeqID 2182	SA-820.1	Contig123 (27245-28117 m)	64	Identities = 113/278 (40%), Positives = 192/278 (68%), Gaps = 1/278 (0%) dbj BAB05397.1 (AP001512) unknown conserved protein [Bacillus halodurans] Length = 290
SeqID 2183	SA-821.1	Contig123 (26910-27218 p)	43	Identities = 27/88 (30%), Positives = 48/88 (53%), Gaps = 1/88 (1%) dbj BAA11330.1 (D78257) BacB [Enterococcus faecalis] Length = 94
SeqID 2184	SA-822.1	Contig123 (25131-26822 p)	80	Identities = 377/566 (66%), Positives = 464/566 (81%), Gaps = 5/566 (0%) gb AAF86984.1 (AF282249_3 (AF282249) ArgS [Lactococcus lactis subsp. lactis] Length = 564
SeqID 2185	SA-823.1	Contig123 (24472-24918 m)	58	Identities = 62/139 (44%), Positives = 89/139 (63%) sp Q54870 ARGR_STRPN_ARGININE_REPRESSOR pir B28667 hypothetical protein C - Streptococcus pneumoniae gb AAA88596.1 (M18729) unknown protein [Streptococcus pneumoniae] Length = 148
SeqID 2186	SA-825.1	Contig123 (21839-24415 m)	81	Identities = 593/858 (69%), Positives = 698/858 (81%), Gaps = 14/858 (1%) sp P10564 HEXA_STRPN_DNA_MISMATCH_REPAIR_PROTEIN_HEXA pir C28667 DNA mismatch repair protein hexA - Streptococcus pneumoniae gb AAA88597.1 (M18729) mismatch repair protein [Streptococcus pneumoniae] Length = 844
SeqID 2187	SA-826.1	Contig123 (21579-21782 p)	76	Identities = 48/63 (76%), Positives = 56/63 (88%) emb CAA71254.1 (Y10181) cold shock protein [Lactococcus lactis] emb CAA76695.1 (Y17215) cold shock protein B [Lactococcus lactis] Length = 66

SeqID 2188	SA-827.1	Contig123 (19380-21359 m)	79	Identities = 452/657 (68%), Positives = 543/657 (81%), Gaps = 8/657 (1%) sp P14160 HEXB_STRPN DNA MISMATCH REPAIR PROTEIN HEXB pir A33589 mismatch repair protein hexB [validated] - Streptococcus pneumoniae gb AAA88600.1 (M29686) mismatch repair protein [Streptococcus pneumoniae] Length = 649
SeqID 2189	SA-828.1	Contig123 (18116-19348 m)	55	Identities = 145/401 (36%), Positives = 236/401 (58%), Gaps = 4/401 (0%) pir S58131 integral membrane protein LmrP - Lactococcus lactis emb CAA61918.1 (X89779) LmrP integral membrane protein [Lactococcus lactis] Length = 408
SeqID 2190	SA-829.1	Contig123 (17524-18114 m)	61	Identities = 86/201 (42%), Positives = 122/201 (59%), Gaps = 6/201 (2%) dbj BAB04943.1 (AP001511) holliday junction DNA helicase [Bacillus halodurans] Length = 203
SeqID 2191	SA-83.1	Contig131 (34805-35092 p)	44	Identities = 24/85 (28%), Positives = 45/85 (52%), Gaps = 8/85 (9%) sp P37466 VEG_BACSU VEG PROTEIN pir S66073 veg protein - Bacillus subtilis dbj BAA05279.1 (D26185) unknown [Bacillus subtilis] emb CAB11820.1 (Z99104) veg [Bacillus subtilis] Length = 86
SeqID 2192	SA-831.1	Contig123 (16950-17501 m)	63	Identities = 87/176 (49%), Positives = 122/176 (68%), Gaps = 1/176 (0%) sp P05100 3MG1_ECOLI DNA-3-METHYLADENINE GLYCOSYLASE I (3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I) (DNA-3-METHYLADENINE GLYCOSIDASE I) pir DGE1M1 3-methyladenine-DNA glycosylase (EC 3.2.2.-) I - Escherichia coli emb CAA27472.1 (X03845) TAGI (aa 1-187) [Escherichia coli] gb AA24658.1 (J02606) 3-methyladenine-DNA glycosylase I (tag) [Escherichia coli] gb AAB18526.1 (U00039) 3-methyladenine DNA glycosylase I, constitutive [Escherichia coli] gb AAC76573.1 (AE000432) 3-methyl-adenine DNA glycosylase I, constitutive [Escherichia coli K12] Length = 187

SeqID 2193	SA-832.1	Contig123 (15602-16861 m)	80	Identities = 292/417 (70%), Positives = 343/417 (82%), Gaps = 1/417 (0%) sp P54184 CINA_STRPN PUTATIVE COMPETENCE DAMAGE PROTEIN (EXPORTED PROTEIN 10) emb CAA84071.1 (Z34303) ClnA protein [Streptococcus pneumoniae] Length = 418
SeqID 2194	SA-833.1	Contig123 (14389-15528 m)	88	Identities = 340/340 (100%), Positives = 340/340 (100%) gb AAG30537.1 AF307982_1 (AF307982) RecA [Streptococcus agalactiae] Length = 340
SeqID 2195	SA-834.1	Contig123 (13775-14185 m)	67	Identities = 61/127 (48%), Positives = 93/127 (73%) gb AAF21893.1 AF103794_1 (AF103794) unknown [Listeria monocytogenes] Length = 131
SeqID 2196	SA-835.1	Contig123 (13307-13573 m)	70	Identities = 50/90 (55%), Positives = 67/90 (73%), Gaps = 1/90 (1%) dbj BAB04987.1 (AP001511) unknown [Bacillus halodurans] Length = 90
SeqID 2197	SA-837.1	Contig123 (12888-13298 m)	68	Identities = 82/138 (59%), Positives = 109/138 (78%), Gaps = 1/138 (0%) sp O34634 YRKR_BACSU HYPOTHETICAL 15.2 KDA PROTEIN IN UDK-ALAS INTERGENIC REGION pir D69979 conserved hypothetical protein yrkK - Bacillus subtilis emb CAB14681.1 (Z99117) similar to hypothetical proteins [Bacillus subtilis] emb CAB14698.1 (Z99118) similar to hypothetical proteins [Bacillus subtilis] Length = 138
SeqID 2198	SA-838.1	Contig123 (12545-12862 m)	60	Identities = 50/95 (52%), Positives = 66/95 (68%), Gaps = 5/95 (5%) pir A69982 hypothetical protein yrzB - Bacillus subtilis emb CAB14680.1 (Z99117) yrzB [Bacillus subtilis] emb CAB14697.1 (Z99118) yrzB [Bacillus subtilis] Length = 93
SeqID 2199	SA-839.2	Contig123 (10754-12295 m)	14	Identities = 39/172 (22%), Positives = 75/172 (42%), Gaps = 9/172 (5%) ref NP_066532.1 ABC transporter subunit [Naegleria gruberi] gb AAG17810.1 AF288092_35 (AF288092) ABC transporter subunit [Naegleria gruberi] Length = 210

SeqID 2200	SA-84.1	Contig131 (33438-34793 p)	95	Identities = 398/452 (88%), Positives = 432/452 (95%), Gaps = 1/452 (0%) gb AAF98347.1 (AF280763) DNA polymerase III delta prime subunit [Streptococcus pyogenes] Length = 455
SeqID 2201	SA-842.2	Contig112 (127-1845 p)	45	Identities = 193/426 (45%), Positives = 278/426 (64%), Gaps = 28/426 (6%) pir B82263 PTS system, trehalose-specific IIBC component VC0910 [imported] - Vibrio cholerae (group O1 strain N16961) gb AAF94072.1 (AE004175) PTS system, trehalose-specific IIBC component [Vibrio cholerae] Length = 478
SeqID 2202	SA-843.1	Contig112 (2067-3692 p)	80	Identities = 380/547 (69%), Positives = 437/547 (79%), Gaps = 13/547 (2%) gb AAB65079.1 (U35633) dextran glucosidase DexS [Streptococcus suis] Length = 542
SeqID 2203	SA-844.1	Contig112 (3912-5948 p)	53	Identities = 203/697 (29%), Positives = 362/697 (51%), Gaps = 34/697 (4%) dbj BAB03939.1 (AP001507) unknown conserved protein [Bacillus halodurans] Length = 701
SeqID 2204	SA-845.1	Contig112 (5951-6235 p)	50	Identities = 26/82 (31%), Positives = 52/82 (62%), Gaps = 2/82 (2%) dbj BAB03941.1 (AP001507) unknown conserved protein [Bacillus halodurans] Length = 89
SeqID 2205	SA-846.1	Contig112 (6248-7603 p)	64	Identities = 203/453 (44%), Positives = 293/453 (63%), Gaps = 22/453 (4%) pir T37066 probable integral membrane protein - Streptomyces coelicolor emb CAB52363.1 (AL109747) putative integral membrane protein [Streptomyces coelicolor A3(2)] Length = 516
SeqID 2206	SA-847.1	Contig112 (7542-7700 m)	No Hits found	
SeqID 2207	SA-848.1	Contig112 (7606-8463 p)	54	Identities = 105/269 (39%), Positives = 158/269 (58%), Gaps = 4/269 (1%) sp Q58094 TKTN_METJA PUTATIVE TRANSKETOLASE N-TERMINAL SECTION (TK) pir A64385 transketolase (EC 2.2.1.1) - Methanococcus jannaschii gb AAB98676.1 (U67515) transketolase [Methanococcus jannaschii] Length = 274

SeqID 2208	SA-849.1	Contig112 (8460-9389 p)	50	<p>Identities = 99/310 (31%), Positives = 172/310 (54%), Gaps = 7/310 (2%) sp Q58092 TKTC_METJA PUTATIVE TRANSKETOLASE C-TERMINAL SECTION (TK) pir G64384 transketolase (EC 2.2.1.1) - Methanococcus jannaschii gb AAB98674.1 (U67515) transketolase [Methanococcus jannaschii] Length = 316</p>
SeqID 2209	SA-85.1	Contig131 (32943-33395 p)	67	<p>Identities = 80/149 (53%), Positives = 105/149 (69%), Gaps = 2/149 (1%) sp P02417 RL9_BACST 50S RIBOSOMAL PROTEIN L9 (BL17) pir R5BS7F ribosomal protein L9 - Bacillus stearothermophilus pdb 487DJK Chain K, Seven Ribosomal Proteins Fitted To A Cryo-Electron Microscopic Map Of The Large 50s Subunit At 7.5 Angstroms Resolution pdb 1DIV Ribosomal Protein L9 Length = 149</p>
SeqID 2210	SA-850.1	Contig112 (9498-10757 p)	42	<p>Identities = 111/438 (25%), Positives = 195/438 (44%), Gaps = 51/438 (11%) pir G83576 hypothetical protein PA0545 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG03934.1 AE004491_1 (AE004491) hypothetical protein [Pseudomonas aeruginosa] Length = 434</p>
SeqID 2211	SA-851.1	Contig112 (10845-11114 p)	70	<p>Identities = 55/89 (61%), Positives = 71/89 (78%) sp P21473 RS15_BACSU 30S RIBOSOMAL PROTEIN S15 (BS18) pir F69700 ribosomal protein S15 (rpsO) - Bacillus subtilis emb CAB02560.1 (Z80835) ribosomal protein S15 [Bacillus subtilis] emb CAB13541.1 (Z99112) ribosomal protein S15 (BS18) [Bacillus subtilis] Length = 89</p>

SeqID 2212	SA-852.1	Contig112 (11495-13624 p)	74	<p>Identities = 431/708 (60%), Positives = 538/708 (75%), Gaps = 8/708 (1%) sp P50849 PNPA_BACSU POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (POLYNUCLEOTIDE PHOSPHORYLASE) (PNPASE) (VEGETATIVE PROTEIN 15) (VEG15) pir S70691 polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) alpha chain pnpA - Bacillus subtilis gb AAC43595.1 (U29668) polynucleotide phosphorylase [Bacillus subtilis] emb CAB13542.1 (Z99112) polynucleotide phosphorylase (PNPase) [Bacillus subtilis] prf 2210369A polynucleotide phosphorylase [Bacillus subtilis] Length = 705</p>
SeqID 2213	SA-853.1	Contig112 (13626-14378 p)	No Hits found	
SeqID 2214	SA-854.1	Contig112 (14387-14971 p)	61	<p>Identities = 92/169 (54%), Positives = 125/169 (73%) emb CAB71304.1 (AJ130879) serine acetyltransferase [Clostridium sticklandii] Length = 191</p>
SeqID 2215	SA-855.1	Contig112 (14972-15163 p)	No Hits found	
SeqID 2216	SA-856.1	Contig112 (15232-16503 p)	70	<p>Identities = 246/467 (52%), Positives = 323/467 (68%), Gaps = 23/467 (4%) sp Q06752 SYC_BACSU CYSTEINYL-TRNA SYNTHETASE (CYSTEINE--TRNA LIGASE) (CYSRS) pir C53402 cysteine--tRNA ligase (EC 6.1.1.16) cysS - Bacillus subtilis gb AAA21798.1 (L14580) cysteinyl-tRNA synthetase [Bacillus subtilis] dbj BAA05328.1 (D26185) cysteinyl-tRNA synthetase [Bacillus subtilis] emb CAA52167.1 (X73989) cysteine--tRNA ligase [Bacillus subtilis] emb CAB11870.1 (Z99104) cysteinyl-tRNA synthetase [Bacillus subtilis] Length = 466</p>
SeqID 2217	SA-857.1	Contig112 (16496-16882 p)	65	<p>Identities = 58/122 (47%), Positives = 87/122 (70%) pir C69742 conserved hypothetical protein yzC - Bacillus subtilis emb CAB11871.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 143</p>

SeqID 2218	SA-858.1	Contig112 (16985-17740 p)	64	Identities = 113/244 (46%), Positives = 163/244 (66%), Gaps = 6/244 (2%) sp Q06753 YACO_BACSU_HYPOTHETICAL TRNA/RRNA METHYLTRANSFERASE YACO pir S66124 conserved hypothetical protein yacO - Bacillus subtilis dbj BAA05329.1 (D26185) unknown.[Bacillus subtilis] emb CAB11872.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 249
SeqID 2219	SA-859.1	Contig112 (17737-18255 p)	52	Identities = 66/147 (44%), Positives = 93/147 (62%), Gaps = 2/147 (1%) sp P37574 YACP_BACSU_HYPOTHETICAL 19.7 KD PROTEIN IN CYSS 3 REGION pir S66125 conserved hypothetical protein yacP - Bacillus subtilis dbj BAA05330.1 (D26185) unknown [Bacillus subtilis] emb CAB11873.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 170
SeqID 2220	SA-86.1	Contig131 (31063-32937 p)	56	Identities = 212/660 (32%), Positives = 376/660 (56%), Gaps = 14/660 (2%) dbj BAB07750.1 (AP001520) unknown conserved protein in B. subtilis [Bacillus halodurans] Length = 654
SeqID 2221	SA-861.1	Contig112 (18348-19208 p)	53	Identities = 100/284 (35%), Positives = 157/284 (55%), Gaps = 6/284 (2%) pir D69841 hypothetical protein yitS - Bacillus subtilis emb CAB01833.1 (Z79580) putative ORF [Bacillus subtilis] emb CAA70629.1 (Y09476) YitS [Bacillus subtilis] emb CAB12951.1 (Z99109) yitS [Bacillus subtilis] Length = 283
SeqID 2222	SA-863.2	Contig112 (20193-21194 m)	35	Identities = 73/255 (28%), Positives = 117/255 (45%), Gaps = 18/255 (7%) emb CAA50689.1 (X71844) putative transposase [Clostridium perfringens] Length = 350
SeqID 2223	SA-864.3	Contig127 (33995-35254 p)	84	Identities = 300/419 (71%), Positives = 359/419 (85%), Gaps = 1/419 (0%) gb AAK04483.1 AE006275_7 (AE006275) tyrosyl-tRNA synthetase 1 (EC 6.1.1.1) [Lactococcus lactis subsp. lactis] Length = 419

SeqID 2224	SA-866.2	Contig127 (31587-33884 m)	75	Identities = 445/769 (57%), Positives = 582/769 (74%), Gaps = 9/769 (1%) emb CAA05302.1 (AJ002291) penicillin-binding protein 1b [Streptococcus pneumoniae] gb AAF04736.1 AF101781_1 (AF101781) penicillin-binding protein 1b [Streptococcus pneumoniae] Length = 821
SeqID 2225	SA-867.1	Contig127 (31366-31455 m)	No Hits found	Identities = 814/1173 (69%), Positives = 978/1173 (82%), Gaps = 17/1173 (1%) emb CAB56706.1 (Y16468) DNA-dependent RNA polymerase subunit beta [Listeria monocytogenes] Length = 1184
SeqID 2226	SA-868.1	Contig127 (27488-31063 m)	81	
SeqID 2227	SA-87.1	Contig131 (28966-30888 p)	78	Identities = 402/627 (64%), Positives = 503/627 (80%), Gaps = 5/627 (0%) db BAB07780.1 (AP001520) glucose-inhibited division protein [Bacillus halodurans] Length = 632
SeqID 2228	SA-872.1	Contig127 (23721-27371 m)	78	Identities = 935/990 (94%), Positives = 959/990 (96%), Gaps = 1/990 (0%) sp P95816 RPOC_STRPY DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) emb CAA65249.1 (X96385) DNA-dependent RNA polymerase [Streptococcus pyogenes] Length = 989
SeqID 2229	SA-874.1	Contig127 (23242-23607 m)	59	Identities = 42/99 (42%), Positives = 75/99 (75%) gb AAC45309.1 (U81957) putative DNA binding protein [Streptococcus gordonii] Length = 122
SeqID 2230	SA-875.1	Contig127 (22098-23069 m)	74	Identities = 202/319 (63%), Positives = 254/319 (79%), Gaps = 1/319 (0%) gb AAC45310.1 (U81957) putative ABC transporter subunit ComYA [Streptococcus gordonii] Length = 319
SeqID 2231	SA-876.1	Contig127 (21161-22252 m)	59	Identities = 161/280 (57%), Positives = 218/280 (77%) gb AAC45311.1 (U81957) putative ABC transporter subunit ComYB [Streptococcus gordonii] Length = 282
SeqID 2232	SA-877.1	Contig127 (20835-21164 m)	58	Identities = 63/103 (61%), Positives = 85/103 (82%) gb AAC45312.1 (U81957) ComYC [Streptococcus gordonii] Length = 105

SeqID 2233	SA-878.1	Contig127 (20447-20860 m)	57	Identities = 54/127 (42%), Positives = 83/127 (64%) gb AAC23740.1 (AF052207) competence protein [Streptococcus pneumoniae] Length = 134
SeqID 2234	SA-879.1	Contig127 (20176-20421 m)	49	Identities = 27/91 (29%), Positives = 50/91 (54%) emb CAA75316.1 (Y15043) hypothetical protein [Lactococcus lactis subsp. cremoris] Length = 128
SeqID 2235	SA-880.1	Contig127 (19761-20222 m)	56	Identities = 61/147 (41%), Positives = 98/147 (66%), Gaps = 3/147 (2%) emb CAA75315.1 (Y15043) homology to ComYD from Streptococcus gordonii, and ComGD from Bacillus subtilis [Lactococcus lactis subsp. cremoris] Length = 150
SeqID 2236	SA-881.1	Contig127 (19412-19783 m)	No Hits found	
SeqID 2237	SA-882.1	Contig127 (18323-19297 m)	53	Identities = 105/329 (31%), Positives = 177/329 (52%), Gaps = 17/329 (5%) sp P37876 YTXK_BACSU HYPOTHETICAL 37.4 KD PROTEIN IN ACKA-SSPA INTERGENIC REGION pir G70003 hypothetical protein ytxK - Bacillus subtilis gb AAC00317.1 (AF008220) YtxK [Bacillus subtilis] emb CAB14926.1 (Z99119) alternate gene name: ythl [Bacillus subtilis] Length = 329
SeqID 2238	SA-883.1	Contig127 (17098-18291 m)	73	Identities = 223/395 (56%), Positives = 293/395 (73%), Gaps = 3/395 (0%) sp P37877 ACKA_BACSU ACETATE KINASE (ACETOKINASE) pir B49935 acetate kinase (EC 2.7.2.1) ackA - Bacillus subtilis gb AAC36857.1 (L17320) acetate kinase [Bacillus subtilis] gb AAC00318.1 (AF008220) acetate kinase [Bacillus subtilis] emb CAB14925.1 (Z99119) acetate kinase [Bacillus subtilis] Length = 395
SeqID 2239	SA-884.1	Contig127 (16741-16947 m)	70	Identities = 39/64 (60%), Positives = 49/64 (75%) pir G75183 probable repressor protein PAB7155 - Pyrococcus abyssi (strain Orsay) emb CAB49550.1 (AJ248284) repressor protein, putative [Pyrococcus abyssi] Length = 73
SeqID 2240	SA-885.2	Contig127 (16049-16510 m)	No Hits found	
SeqID 2241	SA-887.2	Contig139 (52447-52644 p)	No Hits found	

SeqID 2242	SA-889.1	Contig139 (52688-53620 m)	82	<p>Identities = 227/311 (72%), Positives = 269/311 (85%) sp Q9X9S0 PYRD_STRPN DIHYDROOROTATE DEHYDROGENASE (DIHYDROOROTATE OXIDASE) (DHODEHASE) (DHODASE) (DHOD) emb CAB51330.1 (AJ131985) dihydroorotate dehydrogenase [Streptococcus pneumoniae] Length = 311</p>
SeqID 2243	SA-89.1	Contig131 (28206-28817 p)	57	<p>Identities = 71/202 (35%), Positives = 120/202 (59%), Gaps = 5/202 (2%) sp O32244 YVBG_BACSU HYPOTHETICAL 22.6 KD PROTEIN IN OPUCA-ENO INTERGENIC REGION pir F70029 conserved hypothetical protein yvbG - Bacillus subtilis emb CAB15390.1 (Z99121) similar to hypothetical proteins [Bacillus subtilis] Length = 211</p>
SeqID 2244	SA-890.1	Contig139 (53807-55042 m)	72	<p>Identities = 237/410 (57%), Positives = 304/410 (73%) emb CAB89121.1 (AJ277485) beta-lactam resistance factor [Streptococcus pneumoniae] Length = 410</p>
SeqID 2245	SA-891.1	Contig139 (55061-56272 m)	70	<p>Identities = 204/410 (49%), Positives = 286/410 (69%), Gaps = 17/410 (4%) emb CAB89537.1 (AJ250766) MurM protein [Streptococcus pneumoniae] emb CAB89539.1 (AJ250767) MurM protein [Streptococcus pneumoniae] Length = 406</p>
SeqID 2246	SA-892.1	Contig139 (56285-57505 m)	60	<p>Identities = 166/410 (40%), Positives = 250/410 (60%), Gaps = 10/410 (2%) emb CAB89120.1 (AJ277484) beta-lactam resistance factor [Streptococcus pneumoniae] Length = 406</p>
SeqID 2247	SA-893.1	Contig139 (57505-58329 m)	68	<p>Identities = 127/269 (47%), Positives = 189/269 (70%), Gaps = 1/269 (0%) sp P09997 YIDA_ECOLI HYPOTHETICAL 29.7 KD PROTEIN IN IBPA-GYRB INTERGENIC REGION pir QQECGB hypothetical 29.7K protein, ibpA-gyrB intergenic region - Escherichia coli (strain K-12) gb AAC76720.1 (AE000446) orf, hypothetical protein [Escherichia coli K12] Length = 270</p>
SeqID 2248	SA-894.1	Contig139 (58388-59704 m)	61	<p>Identities = 212/429 (49%), Positives = 287/429 (66%), Gaps = 5/429 (1%) dbj BAB07537.1 (AP001520) unknown conserved protein [Bacillus halodurans] Length = 432</p>

SeqID 2249	SA-895.1	Contig139 (59768-60166 p)	No Hits found	Identities = 378/886 (42%), Positives = 542/886 (60%), Gaps = 45/886 (5%) pir H69877 calcium-transporting ATPase homolog yloB - Bacillus subtilis emb CAA74269.1 (Y13937) putative PacL protein [Bacillus subtilis] emb CAB13439.1 (Z99112) similar to calcium-transporting ATPase [Bacillus subtilis] Length = 890
SeqID 2250	SA-896.1	Contig139 (60534-63194 p)	57	
SeqID 2251	SA-897.1	Contig139 (63239-64099 m)	33	Identities = 60/221 (27%), Positives = 100/221 (45%), Gaps = 37/221 (16%) pir E75186 hypothetical protein PAB0011 - Pyrococcus abyssi (strain Orsay) emb CAB48940.1 (AJ248283) hypothetical protein [Pyrococcus abyssi] Length = 248
SeqID 2252	SA-898.1	Contig139 (64251-66182 p)	69	Identities = 315/644 (48%), Positives = 447/644 (68%), Gaps = 7/644 (1%) pir C69621 fructose-bisphosphatase (EC 3.1.3.11) - Bacillus subtilis dbj BAA11277.1 (D78193) yydE [Bacillus subtilis] emb CAB16056.1 (Z99124) fructose-1,6-bisphosphatase [Bacillus subtilis] Length = 671
SeqID 2253	SA-899.1	Contig139 (66272-67396 p)	64	Identities = 177/371 (47%), Positives = 242/371 (64%), Gaps = 6/371 (1%) pir E69820 conserved hypothetical protein yhbA - Bacillus subtilis emb CAB07527.1 (Z93102) hypothetical 48.5 kd protein [Bacillus subtilis] emb CAB12719.1 (Z99108) alternate gene name: ygaP~similar to hypothetical proteins [Bacillus subtilis] Length = 435
SeqID 2254	SA-9.1	Contig137 (36200-36682 p)	No Hits found	
SeqID 2255	SA-90.1	Contig131 (27053-28174 p)	72	Identities = 250/359 (69%), Positives = 292/359 (80%), Gaps = 6/359 (1%) dbj BAB04980.1 (AP001511) (5-methylaminomethyl-2-thiouridylate)-methyltran sferase [Bacillus halodurans] Length = 371

SeqID 2256	SA-901.1	Contig139 (67583-68563 p)	73	Identities = 197/344 (57%), Positives = 255/344 (73%), Gaps = 2/344 (0%) sp P28367 RF2_BACSU PEPTIDE CHAIN RELEASE FACTOR 2 (RF-2) pir JN0146 translation releasing factor RF-2 - Bacillus subtilis gb AAC97534.1 (AF013188) release factor 2 [Bacillus subtilis] gb AAC67303.1 (AF017113) putative peptide chain release factor RF-2 [Bacillus subtilis] emb CAB15546.1 (Z99122) peptide chain release factor 2 [Bacillus subtilis] Length = 366
SeqID 2257	SA-902.1	Contig139 (68582-69274 p)	76	Identities = 138/228 (60%), Positives = 179/228 (77%) pir D69627 cell-division ATP-binding proteinftsE - Bacillus subtilis gb AAC67262.1 (AF017113) cell division ATP-binding protein [Bacillus subtilis] emb CAB15543.1 (Z99122) cell-division ATP-binding protein [Bacillus subtilis] Length = 228
SeqID 2258	SA-904.1	Contig139 (69258-70187 p)	52	Identities = 111/311 (35%), Positives = 181/311 (57%), Gaps = 31/311 (9%) sp O34876 FTSX_BACSU CELL DIVISION PROTEIN FTSX HOMOLOG pir G69627 cell-division proteinftsX - Bacillus subtilis gb AAC67264.1 (AF017113) cell division protein [Bacillus subtilis] emb CAB15542.1 (Z99122) cell-division protein [Bacillus subtilis] Length = 296
SeqID 2259	SA-905.1	Contig139 (70240-70950 m)	41	Identities = 65/185 (35%), Positives = 100/185 (53%), Gaps = 3/185 (1%) pir A75409 carboxymethylenebutenolidase-related protein - Deinococcus radiodurans (strain R1) gb AAF10898.1 AE001979_4 (AE001979) carboxymethylenebutenolidase-related protein [Deinococcus radiodurans] Length = 292
SeqID 2260	SA-906.2	Contig139 (70947-71582 m)	48	Identities = 83/197 (42%), Positives = 114/197 (57%), Gaps = 4/197 (2%) dbj BAB06539.1 (AP001516) BH2820~unknown conserved protein [Bacillus halodurans] Length = 211

SeqID 2261	SA-907.2	Contig130 (21596-21796 p)	76	Identities = 56/66 (84%), Positives = 60/66 (90%) -dbj BAA23749.1 (AB009314) proton-translocating ATPase, c subunit [Streptococcus bovis] Length = 67
SeqID 2262	SA-908.2	Contig130 (21829-22545 p)	75	Identities = 147/238 (61%), Positives = 180/238 (74%) dbj BAA23750.1 (AB009314) proton-translocating ATPase, a subunit [Streptococcus bovis] Length = 239
SeqID 2263	SA-909.1	Contig130 (22563-23060 p)	73	Identities = 103/165 (62%), Positives = 130/165 (78%) sp P95785 ATPF_STRMU ATP SYNTHASE B CHAIN pir JC5737 membrane-bound proton-translocating ATPase (EC 3.6.1.-) b chain - Streptococcus mutans gb AAD13379.1 (U31170) ATPase, b subunit [Streptococcus mutans] Length = 165
SeqID 2264	SA-91.1	Contig131 (26139-26807 m)	64	Identities = 101/216 (46%), Positives = 156/216 (71%), Gaps = 2/216 (0%) dbj BAB06216.1 (AP001515) L-serine dehydratase beta subunit [Bacillus halodurans] Length = 220
SeqID 2265	SA-910.1	Contig130 (23060-23596 p)	68	Identities = 98/178 (55%), Positives = 127/178 (71%) sp O50156 ATPD_STRBO ATP SYNTHASE DELTA CHAIN dbj BAA23752.1 (AB009314) proton-translocating ATPase, delta subunit [Streptococcus bovis] Length = 178
SeqID 2266	SA-911.1	Contig130 (23612-25117 p)	98	Identities = 480/501 (95%), Positives = 497/501 (98%) dbj BAA23753.1 (AB009314) proton-translocating ATPase, alpha subunit [Streptococcus bovis] Length = 501
SeqID 2267	SA-912.1	Contig130 (25133-26014 p)	93	Identities = 254/293 (86%), Positives = 278/293 (94%), Gaps = 2/293 (0%) dbj BAA23754.1 (AB009314) proton-translocating ATPase, gamma subunit [Streptococcus bovis] Length = 291
SeqID 2268	SA-913.1	Contig130 (26088-27494 p)	92	Identities = 434/466 (93%), Positives = 454/466 (97%) sp P95789 ATPB_STRMU ATP SYNTHASE BETA CHAIN gb AAD13383.1 (U31170) ATPase, beta subunit [Streptococcus mutans] Length = 468

SeqID 2269	SA-914.1	Contig130 (27507-27920 p)	85	<p>Identities = 102/138 (73%), Positives = 121/138 (86%), Gaps = 1/138 (0%) dbj BAA23756.1 (AB009314) proton-translocating ATPase, epsilon subunit [Streptococcus bovis]</p> <p>Length = 138</p>
SeqID 2270	SA-916.1	Contig130 (28278-29549 p)	73	<p>Identities = 248/423 (58%), Positives = 323/423 (75%), Gaps = 5/423 (1%) sp P70965 MUA1_BACSU UDP-N-ACETYLGLUCOSAMINE 1-CARBOXYVINYLTRANSFERASE 1 (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT) pir A69662 UDP-N-acetylglucosamine 1-carboxyvinyltransferase murA - Bacillus subtilis emb CAB03688.1 (Z81356) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis] emb CAB15693.1 (Z99122) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis] Length = 436</p>
SeqID 2271	SA-917.1	Contig130 (29817-30674 p)	69	<p>Identities = 167/287 (58%), Positives = 200/287 (69%), Gaps = 15/287 (5%) sp Q03158 NUCE_STRPN DNA-ENTRY NUCLEASE (COMPETENCE-SPECIFIC NUCLEASE) pir S10641 endA protein - Streptococcus pneumoniae emb CAA38134.1 (X54225) membrane nuclease [Streptococcus pneumoniae] Length = 274</p>
SeqID 2272	SA-918.1	Contig130 (30965-32005 p)	77	<p>Identities = 211/341 (61%), Positives = 272/341 (78%), Gaps = 2/341 (0%) sp P17921 SYFA_BACSU PHENYLALANINE-TRNA SYNTHETASE ALPHA CHAIN (PHENYLALANINE-TRNA LIGASE ALPHA CHAIN) (PHERS) pir YFB5A phenylalanine--trna ligase (EC 6.1.1.20) alpha chain - Bacillus subtilis emb CAA99603.1 (Z75208) phenylalanyl-tRNA synthetase beta subunit [Bacillus subtilis] emb CAB14824.1 (Z99118) phenylalanyl-tRNA synthetase (alpha subunit) [Bacillus subtilis] Length = 344</p>

SeqID 2273	SA-919.1	Contig130 (32088-32609 p)	51	<p>Identities = 61/169 (36%), Positives = 100/169 (59%) sp P21340 PAIA_BACSU PROTEASE SYNTHASE AND SPORULATION NEGATIVE REGULATORY PROTEIN PAI 1 emb CAB15205.1 (Z99120) transcriptional regulator [Bacillus subtilis] Length = 172</p>
SeqID 2274	SA-920.1	Contig130 (32663-35068 p)	64	<p>Identities = 376/805 (46%), Positives = 523/805 (64%), Gaps = 6/805 (0%) sp P17922 SYFB_BACSU PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PHERS) pir YFBSB phenylalanine--trna ligase (EC 6.1.1.20) beta chain - Bacillus subtilis emb CAA99564.1 (Z75208) phenylalanyl-tRNA synthetase beta subunit [Bacillus subtilis] emb CAB14823.1 (Z99118) phenylalanyl-tRNA synthetase (beta subunit) [Bacillus subtilis] Length = 804</p>
SeqID 2275	SA-921.1	Contig130 (35137-35805 m)	63	<p>Identities = 105/196 (53%), Positives = 141/196 (71%), Gaps = 1/196 (0%) pir B65023 hypothetical protein b2475 - Escherichia coli (strain K-12) gb AAC75528.1 (AE000334) orf, hypothetical protein [Escherichia coli K12] Length = 287</p>
SeqID 2276	SA-923.1	Contig130 (35916-39149 p)	56	<p>Identities = 363/1093 (33%), Positives = 606/1093 (55%), Gaps = 67/1093 (6%) pir T30307 rexB protein - Lactococcus lactis gb AAC12965.1 (U76424) exonuclease RexB [Lactococcus lactis] Length = 1099</p>
SeqID 2277	SA-925.1	Contig130 (39136-42762 p)	61	<p>Identities = 518/1212 (42%), Positives = 744/1212 (60%), Gaps = 73/1212 (6%) pir T30308 rexA protein - Lactococcus lactis gb AAC12966.1 (U76424) exonuclease RexA [Lactococcus lactis] Length = 1173</p>
SeqID 2278	SA-926.1	Contig130 (42775-43701 p)	47	<p>Identities = 88/306 (28%), Positives = 149/306 (47%), Gaps = 17/306 (5%) ref NP_053049.1 hypothetical protein [Plasmid pNZ4000] gb AAD40365.1 (AF036485) hypothetical protein [Plasmid pNZ4000] Length = 302</p>

SeqID 2279	SA-928.2	Contig130 (43676-45052 m)	95	Identities = 443/462 (95%), Positives = 447/462 (95%), Gaps = 4/462 (0%) emb CAB61255.1 (AJ251564) thiophene degradation protein F [Streptococcus agalactiae] Length = 462
SeqID 2280	SA-93.1	Contig131 (25252-26124 m)	75	Identities = 176/289 (60%), Positives = 224/289 (76%), Gaps = 1/289 (0%) sp O34607 SDHA_BACSU PROBABLE L-SERINE DEHYDRATASE, ALPHA CHAIN (L-SERINE DEAMINASE) (SDH) (L-SD) pir G69879 L-serine dehydratase homolog ylpA - Bacillus subtilis emb CAA74259.1 (Y13937) putative YhaP protein [Bacillus subtilis] emb CAB13459.1 (Z99112) similar to L-serine dehydratase [Bacillus subtilis] Length = 300
SeqID 2281	SA-930.2	Contig117 (7623-8582 m)	76	Identities = 241/316 (76%), Positives = 265/316 (83%) gb AAK05072.1 AE006332_1 (AE006332) ribonucleoside-diphosphate reductase beta chain (EC 1.17.4.1) [Lactococcus lactis subsp. lactis] Length = 325
SeqID 2282	SA-931.2	Contig117 (8785-10944 m)	67	Identities = 367/701 (52%), Positives = 488/701 (69%), Gaps = 19/701 (2%) gb AAD41036.1 AF112535_3 (AF112535) ribonucleotide reductase alpha-chain [Corynebacterium glutamicum] Length = 707
SeqID 2283	SA-932.2	Contig117 (11022-11246 m)	67	Identities = 42/70 (60%), Positives = 53/70 (75%) sp Q48708 NRDH_LACLC GLUTAREDOXIN-LIKE PROTEIN NRDH emb CAA63372.1 (X92690) glutaredoxin-like protein [Lactococcus lactis] Length = 72
SeqID 2284	SA-934.1	Contig117 (11628-11891 p)	92	Identities = 86/87 (98%), Positives = 87/87 (99%) dbj BAA78048.1 (AB027569) histidine containing protein [Streptococcus bovis] Length = 87
SeqID 2285	SA-936.1	Contig117 (11896-13629 p)	95	Identities = 533/577 (92%), Positives = 559/577 (96%) dbj BAA78049.1 (AB027569) phosphoenolpyruvate-protein phosphotransferase [Streptococcus bovis] Length = 577

SeqID 2286	SA-937.1	Contig117 (13779-15206 p)	80	<p>sp Q59931 GAPN_STRMU NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE) (GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE [NADP+]) (TRIOSEPHOSPHATE DEHYDROGENASE) pir A57151 glyceralddehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9) - Streptococcus mutans pdb 1EUH A Chain A, Apo Form Of A NADp Dependent Aldehyde Dehydrogenase From Streptococcus Mutans pdb 1EUH B Chain B, Apo Form Of A NADp Dependent Aldehyde Dehydrogenase From Streptococcus Mutans pdb 1EUH C Chain C, Apo Form Of A NADp Dependent Aldehyde Dehydrogenase From Streptococcus Mutans pdb 1EUH D Chain D, Apo Form Of A NADp Dependent Aldehyde Dehydrogenase From Streptococcus Mutans pdb 2EUH A Chain A, Holo Form Of A NADp Dependent Aldehyde Dehydrogenase Complex With NADp+ pdb 2EUH B Chain B, Holo Form Of A NADp Dependent Aldehyde Dehydrogenase Complex With NADp+ Length = 463</p>
SeqID 2287	SA-939.1	Contig117 (15346-16599 p)	54	<p>Identities = 133/439 (30%), Positives = 232/439 (52%), Gaps = 20/439 (4%) emb CAB96552.1 (AJ251472) peptidoglycan GlcNAc deacetylase [Streptococcus pneumoniae] Length = 463</p>
SeqID 2288	SA-94.1	Contig131 (24440-25117 p)	32	<p>Identities = 70/115 (60%), Positives = 78/115 (66%), Gaps = 6/115 (5%) dbj BAA95959.1 (AB042839) secretory protein SAI-B [Staphylococcus aureus] Length = 233</p>
SeqID 2289	SA-941.1	Contig117 (16630-17712 m)	54	<p>Identities = 138/343 (40%), Positives = 200/343 (58%), Gaps = 9/343 (2%) pir C69813 RNA helicase homolog yfmL - Bacillus subtilis dbj BAA2326.1 (D86417) YfmL [Bacillus subtilis] emb CAB12572.1 (Z99108) similar to RNA helicase [Bacillus subtilis] Length = 376</p>

SeqID 2290	SA-942.1	Contig117 (17857-18486 p)	79	Identities = 132/207 (63%), Positives = 167/207 (79%) sp O32033 URK_BACSU URIDINE KINASE (URIDINE MONOPHOSPHOKINASE) pir G69728 uridine kinase udk - Bacillus subtilis emb CAB14675.1 (Z99117) uridine kinase [Bacillus subtilis] Length = 211
SeqID 2291	SA-943.1	Contig117 (18573-19070 p)	54	Identities = 66/146 (45%), Positives = 95/146 (64%) pir F82706 conserved hypothetical protein XF1230 [imported] - Xylella fastidiosa (strain 9a5c) gb AAF84040.1 AE003957_1 (AE003957) conserved hypothetical protein [Xylella fastidiosa] Length = 162
SeqID 2292	SA-944.1	Contig117 (19070-20734 p)	84	Identities = 408/558 (73%), Positives = 473/558 (84%), Gaps = 6/558 (1%) gb AAF98348.1 (AF280764) DNA polymerase III gamma/tau subunits [Streptococcus pyogenes] Length = 556
SeqID 2293	SA-945.1	Contig117 (20847-21017 p)	No Hits found	
SeqID 2294	SA-946.1	Contig117 (20998-21933 m)	50	Identities = 103/315 (32%), Positives = 169/315 (52%), Gaps = 18/315 (5%) dbj BAB05404.1 (AP001512) transcriptional repressor of the biotin operon [Bacillus halodurans] Length = 323
SeqID 2295	SA-947.1	Contig117 (22118-23314 p)	78	Identities = 265/390 (67%), Positives = 325/390 (82%), Gaps = 1/390 (0%) sp P54419 METHK_BACSU S-ADENOSYLMETHIONINE SYNTHETASE (METHIONINE ADENOSYLTRANSFERASE) (ADOMET SYNTHETASE) pir D69657 methionine adenosyltransferase (EC 2.5.1.6) - Bacillus subtilis gb AAC00242.1 (AF008220) SAM synthase [Bacillus subtilis] emb CAB15033.1 (Z99119) S-adenosylmethionine synthetase [Bacillus subtilis] Length = 400
SeqID 2296	SA-948.1	Contig117 (23809-25740 p)	40	Identities = 135/648 (20%), Positives = 261/648 (39%), Gaps = 90/648 (13%) emb CAB39037.2 (AL034559) hypothetical protein, PFC0960c [Plasmodium falciparum] Length = 1946
SeqID 2297	SA-949.1	Contig117 (25807-26352 p)	No Hits found	
SeqID 2298	SA-950.1	Contig117 (26503-26640 p)	No Hits found	

SeqID 2299	SA-951.1	Contig117 (26745-27314 p)	34	Identities = 50/177 (28%), Positives = 81/177 (45%), Gaps = 31/177 (17%) pir B72471 hypothetical protein APE2411 - Aeropyrum pernix (strain K1) dbj BAA81426.1 (AP000064) 197aa long hypothetical protein [Aeropyrum pernix] Length = 197
SeqID 2300	SA-952.1	Contig117 (27311-27865 p)	63	Identities = 71/182 (39%), Positives = 121/182 (66%), Gaps = 3/182 (1%) pir A69859 hypothetical protein ykoE - Bacillus subtilis emb CAA05602.1 (AJ002571) YkoE [Bacillus subtilis] emb CAB13180.1 (Z99110) ykoE [Bacillus subtilis] Length = 199
SeqID 2301	SA-953.1	Contig117 (27869-29071 p)	55	Identities = 139/450 (30%), Positives = 224/450 (48%), Gaps = 68/450 (15%) pir H69858 cation ABC transporter (ATP-binding protein) homolog ykoD - Bacillus subtilis emb CAA05601.1 (AJ002571) YkoD [Bacillus subtilis] emb CAB13179.1 (Z99110) similar to cation ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 490
SeqID 2302	SA-954.1	Contig107 (489-1790 m)	No Hits found	
SeqID 2303	SA-955.1	Contig107 (1924-2223 m)	No Hits found	
SeqID 2304	SA-956.1	Contig107 (2234-3481 m)	58	Identities = 167/436 (38%), Positives = 248/436 (56%), Gaps = 44/436 (10%) dbj BAB07254.1 (AP001519) unknown [Bacillus halodurans] Length = 449
SeqID 2305	SA-957.1	Contig107 (3478-5109 m)	56	Identities = 194/549 (35%), Positives = 316/549 (57%), Gaps = 36/549 (6%) gb AAB06502.1 (U50902) relaxase [Lactococcus lactis subsp. lactis] Length = 563
SeqID 2306	SA-958.1	Contig107 (5081-5455 m)	40	Identities = 31/102 (30%), Positives = 53/102 (51%), Gaps = 2/102 (1%) gb AAF72356.1 AF192329_17 (AF192329) unknown [Enterococcus faecalis] Length = 109
SeqID 2307	SA-959.1	Contig107 (5458-6021 m)	No Hits found	
SeqID 2308	SA-96.1	Contig131 (23828-24307 p)	36	Identities = 54/98 (55%), Positives = 68/98 (69%), Gaps = 4/98 (4%) emb CAA69725.1 (Y08498) aggregation promoting protein [Lactobacillus gasseri] Length = 297

SeqID 2309	SA-960.1	Contig107 (6024-6353 m)	No Hits found	
SeqID 2310	SA-961.1	Contig107 (6364-6663 m)	No Hits found	
SeqID 2311	SA-964.1	Contig107 (6715-9951 m)	15	Identities = 106/344 (30%), Positives = 175/344 (50%), Gaps = 21/344 (6%) ref NP_047307.1 LtrC [Lactococcus lactis] pir T43094 hypothetical protein - Lactococcus lactis plasmid pMRC01 gb AAC56005.1 (AE001272) LtrC [Lactococcus lactis] Length = 355
SeqID 2312	SA-966.1	Contig107 (9953-10318 m)	No Hits found	
SeqID 2313	SA-967.1	Contig107 (10360-12360 m)	42	Identities = 177/530 (33%), Positives = 296/530 (55%), Gaps = 22/530 (4%) gb AAF72343.1 AF192329_4 (AF192329) TrsK-like protein [Enterococcus faecalis] Length = 564
SeqID 2314	SA-968.1	Contig107 (12405-12896 m)	32	Identities = 37/114 (32%), Positives = 56/114 (48%), Gaps = 10/114 (8%) pir E72215 oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima (strain MSB8) gb AAD36811.1 AE001813_3 (AE001813) oligopeptide ABC transporter, periplasmic oligopeptide-binding protein [Thermotoga maritima] Length = 642
SeqID 2315	SA-969.1	Contig107 (12918-13700 m)	31	Identities = 47/187 (25%), Positives = 84/187 (44%), Gaps = 21/187 (11%) pir S45089 hypothetical protein eta - Streptococcus pyogenes plasmid pDB101 pir S68123 hypothetical protein eta - Streptococcus pyogenes plasmid pMD101 emb CAA47095.1 (X66468) orf eta [Streptococcus pyogenes] Length = 231
SeqID 2316	SA-970.1	Contig107 (13700-13972 m)	No Hits found	
SeqID 2317	SA-971.1	Contig107 (13992-14597 m)	No Hits found	
SeqID 2318	SA-973.1	Contig107 (14618-17308 m)	7	Identities = 54/133 (40%), Positives = 66/133 (49%), Gaps = 9/133 (6%) gb AAC61959.1 (AF051917) putative membrane protein TraG [Staphylococcus aureus] prf 2004267H traG protein [Staphylococcus sp.] Length = 358
SeqID 2319	SA-974.1	Contig107 (17287-17871 p)	No Hits found	

SeqID 2320	SA-975.1	Contig107 (18065-20356 m)	50	Identities = 233/769 (30%), Positives = 387/769 (50%), Gaps = 29/769 (3%) gb AAF72347.1 AF192329_8 (AF192329) TrsE-like protein [Enterococcus faecalis] Length = 799
SeqID 2321	SA-976.2	Contig114 (16174-17361 m)	33	Identities = 77/279 (27%), Positives = 136/279 (48%), Gaps = 14/279 (5%) sp Q54955 CIAH_STRPN SENSOR PROTEIN CIAH pir S49545 histidine kinase - Streptococcus pneumoniae emb CAA54466.1 (X77249) histidine kinase [Streptococcus pneumoniae] emb CAB54565.1 (AJ005926) histidine kinase [Streptococcus pneumoniae] Length = 444
SeqID 2322	SA-977.1	Contig114 (14493-16028 m)	90	Identities = 403/511 (78%), Positives = 466/511 (90%) gb AAC05774.1 (AF051356) D-alanine-D-alanyl carrier protein ligase [Streptococcus mutans] Length = 516
SeqID 2323	SA-978.1	Contig114 (13231-14496 m)	81	Identities = 296/418 (70%), Positives = 351/418 (83%), Gaps = 1/418 (0%) gb AAC05775.1 (AF051356) integral membrane protein [Streptococcus mutans] Length = 420
SeqID 2324	SA-98.1	Contig131 (22758-23552 p)	67	Identities = 133/263 (50%), Positives = 191/263 (72%) pir F69742 hypothetical protein ybaF - Bacillus subtilis emb CAB11923.1 (Z99104) ybaF [Bacillus subtilis] Length = 265
SeqID 2325	SA-980.1	Contig114 (12977-13216 m)	85	Identities = 65/79 (82%), Positives = 74/79 (93%) gb AAC05776.1 (AF051356) D-alanyl carrier protein [Streptococcus mutans] gb AAC29040.1 (AF050517) unknown [Streptococcus mutans] gb AAC06286.1 (AF049357) Glg3 [Streptococcus mutans] Length = 79
SeqID 2326	SA-981.1	Contig114 (11722-12984 m)	72	Identities = 238/421 (56%), Positives = 308/421 (72%), Gaps = 1/421 (0%) gb AAC29041.1 (AF050517) unknown [Streptococcus mutans] Length = 421
SeqID 2327	SA-982.1	Contig114 (11155-11472 m)	No Hits found	
SeqID 2328	SA-983.1	Contig114 (9863-11155 m)	19	Identities = 49/185 (26%), Positives = 85/185 (45%), Gaps = 46/185 (24%) dbj BAB01041.1 (AB022220) gene_id:MLN21.14-unknown protein [Arabidopsis thaliana] Length = 507

SeqID 2329	SA-984.1	Contig114 (9480-9872 m)	No Hits found	
SeqID 2330	SA-985.1	Contig114 (9189-9479 m)	No Hits found	
SeqID 2331	SA-986.2	Contig114 (8225-9007 m)	55	<p>Identities = 115/254 (45%), Positives = 172/254 (67%)</p> <p>sp P37545 YABD_BACSU PUTATIVE DEOXYRIBONUCLEASE YABD pir S66068 conserved hypothetical protein yabD - Bacillus subtilis dbj BAA05274.1 (D26185) unknown [Bacillus subtilis] emb CAB11815.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 255</p>
SeqID 2332	SA-987.2	Contig114 (7672-8241 m)	60	<p>Identities = 82/179 (45%), Positives = 117/179 (64%), Gaps = 4/179 (2%) sp P37547 YABF_BACSU HYPOTHETICAL 20.7 KD PROTEIN IN METS-KSGA INTERGENIC REGION pir S66070 conserved hypothetical protein yabF - Bacillus subtilis dbj BAA05276.1 (D26185) unknown [Bacillus subtilis] emb CAB11817.1 (Z99104) similar to hypothetical proteins [Bacillus subtilis] Length = 186</p>
SeqID 2333	SA-988.2	Contig114 (7219-7668 m)	39	<p>Identities = 39/121 (32%), Positives = 60/121 (49%), Gaps = 12/121 (9%) pir A83524 hypothetical protein PA0959 [imported] - Pseudomonas aeruginosa (strain PAO1) gb AAG04348.1 AE004530_1 (AE004530) hypothetical protein [Pseudomonas aeruginosa] Length = 209</p>

SeqID 2334	SA-989.1	Contig114 (6320-7192 m)	72	<p>Identities = 156/284 (54%), Positives = 215/284 (74%), Gaps = 2/284 (0%) sp P37468 KSGA_BACSU DIMETHYLADENOSINE TRANSFERASE (S-ADENOSYLMETHIONINE-6-N, N - ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RNA DIMETHYLASE) (HIGH LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN DIMETHYLTRANSFERASE) pir S66071 probable (adenine-N6,N6-)-dimethyltransferase (EC 2.1.1.-) ksgA - Bacillus subtilis dbj BAA05277.1 (D26185) high level kasgamyacin resistance [Bacillus subtilis] emb CAB11818.1 (Z99104) dimethyladenosine transferase [Bacillus subtilis] Length = 292</p>
SeqID 2335	SA-99.1	Contig131 (21923-22765 p)	63	<p>Identities = 141/242 (58%), Positives = 188/242 (77%), Gaps = 1/242 (0%) pir E69742 ABC transporter (ATP-binding protein) homolog ybaE - Bacillus subtilis dbj BAA10984.1 (D64126) unknown [Bacillus subtilis] emb CAB11922.1 (Z99104) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] Length = 276</p>
SeqID 2336	SA-990.1	Contig114 (5465-6316 m)	51	<p>Identities = 94/278 (33%), Positives = 147/278 (52%), Gaps = 16/278 (5%) sp P42313 YXJB_BACSU HYPOTHETICAL 31.5 KD PROTEIN IN KATB 3 REGION pir H70078 conserved hypothetical protein yxjB - Bacillus subtilis dbj BAA11703.1 (D83026) homologous to SwissProt:YEBH_ECOLI hypothetical protein; hypothetical [Bacillus subtilis] emb CAB15927.1 (Z99123) similar to hypothetical proteins [Bacillus subtilis] emb CAB15937.1 (Z99124) similar to hypothetical proteins [Bacillus subtilis] Length = 282</p>

SeqID 2337	SA-991.1	Contig114 (4437-5309 m)	69	Identities = 149/296 (50%), Positives = 204/296 (68%), Gaps = 14/296 (4%) pir A69879 conserved hypothetical protein yloQ - Bacillus subtilis emb CAA74251.1 (Y13937) YloQ protein [Bacillus subtilis] emb CAB13451.1 (Z99112) similar to hypothetical proteins [Bacillus subtilis] Length = 298
SeqID 2338	SA-993.1	Contig114 (3768-4430 m)	68	Identities = 112/211 (53%), Positives = 152/211 (71%) dbj BAB06221.1 (AP001515) unknown conserved protein [Bacillus halodurans] Length = 216
SeqID 2339	SA-994.1	Contig114 (3143-3775 m)	52	Identities = 75/220 (34%), Positives = 112/220 (50%), Gaps = 18/220 (8%) pir C69879 hypothetical protein yloS - Bacillus subtilis emb CAA74253.1 (Y13937) YloS protein [Bacillus subtilis] emb CAB13453.1 (Z99112) yloS [Bacillus subtilis] Length = 214
SeqID 2340	SA-995.1	Contig114 (1867-3099 m)	54	Identities = 157/393 (39%), Positives = 238/393 (59%), Gaps = 3/393 (0%) pir D82810 conserved hypothetical protein XF0413 [imported] - Xylella fastidiosa (strain 9a5c) gb AAF83223.1 AE003892_8 (AE003892) conserved hypothetical protein [Xylella fastidiosa] Length = 456
SeqID 2341	SA-996.1	Contig114 (936-1877 m)	61	Identities = 139/313 (44%), Positives = 196/313 (62%), Gaps = 6/313 (1%) gb AAC44803.1 (U21636) cmp-binding-factor 1 [Staphylococcus aureus] Length = 313
SeqID 2342	SA-997.1	Contig114 (27-839 m)	71	Identities = 142/270 (52%), Positives = 196/270 (72%), Gaps = 1/270 (0%) emb CAA10902.1 (AJ222642) purR [Lactococcus lactis] Length = 271
SeqID 2343	SA-998.2	Contig127 (15315-15944 m)	30	Identities = 47/120 (39%), Positives = 69/120 (57%), Gaps = 5/120 (4%) dbj BAA11325.1 (D78257) ORF8 [Enterococcus faecalis] Length = 120

SeqID 2344	SA-999.1	Contig127 (14524-15294 p)	59	Identities = 98/265 (36%), Positives = 161/265 (59%), Gaps = 4/265 (1%) sp O66553 PROC_AQUAE PYRROLINE-5-CARBOXYLATE REDUCTASE (P5CR) (P5C REDUCTASE) pir F70315 pyrroline carboxylate reductase - Aquifex aeolicus gb AAC06504.1 (AE000676) pyrroline carboxylate reductase [Aquifex aeolicus] Length = 265
------------	----------	---------------------------	----	---

TABLEAU 2. Protéines de surface de *Streptococcus agalactiae*

	SEQ ID	N° IPF	Motif LPXTG	% identité/ similitude
1	886	18	LPHTG	24/37 : Sec (Surface exclusion protéine) <i>E. faecalis</i>
2	386	1250	LPSTG	22/40 : Sec10 (Surface exclusion protéine) <i>Enterococcus faecalis</i> plasmid pCF10
3	1669	280	LPKTG	50/65 : pullulanase <i>S. pneumoniae</i>
4	1895	523	LPATG	90/91 : alpha like protéine <i>S. agalactiae</i>
5	806	1716	LPSTG	31/46 : M-like protéine <i>S. equi</i>
6	1183	2140	LPLTG	36/60 : cell-division protein homolog ywcF <i>B. subtilis</i> 34/56 : stage V sporulation protein E <i>B. halodurans</i>
7	1346	2337	LPKTG	25/38 : Unknown <i>B. halodurans</i>
8	1942	571	LPSTG	50/60 : Hemagglutinine <i>S. gordonii</i>
9	2129	765	LPNTG	28/43 : sdrc <i>S. aureus</i>
10	2046	678	LPKTG	57/70 : putative cyclo-nucleotide phosphodiesterase <i>Strep. dysgalactiae</i> susp. <i>equismilis</i>
11	618	1503	LPKTG	49/65 : celle envelope proteinase <i>S. thermophilus</i>
12	1227	220	LPSTG	30/43 : SspB (adhesion) <i>S. gordonii</i>
13	1954	584	LPKTG	26/48 : hypothetical serine riche repeat prot <i>S. pombe</i>
14	1493	2495	LPKTG	23/47 : gene drosophile
15	1955	585	LPKTG	30/46 : 6-aminohexanoate-cyclic-dimer hydrolase <i>Deinococcus radiodurans</i> .
16	943	1861	LPKTG	50/62 : hypothetical protein 2 <i>S. mutans</i>
17	1221	2192	LPKTG	35/48 : <i>C. elegans</i> UNC-89 (6642 aa)
18	613	15	LPSTG	37/52 : SpaA : Ag de surface de <i>S. sobrinus</i>
19	382	1247	LPSTG	38/52 : SpaA : Ag de surface de <i>S. sobrinus</i>
20	792	17	LPKTG	30 /47 : hypothetical protéine iota de <i>S. pyogènes</i> plasmid pDB101
21	383	1248	LPRTG	38/52 : hypothetical protéine iota de <i>S. pyogènes</i> plasmid pDB101
22	1418	2414	LPSTG	33/47 : orfC <i>E. faecalis</i> , plasmid pAM-beta -1
Autres protéines possibles (IPXTG, FPXTG)				
23	2171	807	IPQTG	31/43 : Inconnue <i>B. Halodurans</i> 29/41 : fimbrial SU <i>Actinomyces naeslundii</i> 24 :41 : putative cell-surface adhesin SdrF [<i>Staphylococcus epidermidis</i>]
24	1343	2334	IPQTG	27/38 : hypothetical protein 2 <i>L. leichmanii</i> 23/38 : fimbrial SU <i>Actinomyces naeslundii</i>
25	659	1551	FPKTG	37/50 : fibronectin binding protein I <i>S. pyogenes</i>

TABLEAU 3: Liste des phases codantes annotées identifiées par l'analyse de la séquence génomique complète SEQ ID N°2345

SEQ ID (ADN)	SEQ ID (Protéine)	Anciens SEQ ID (Protéine)	Protéine N°	Positions / Sens (p=plus, m=moins)	Fonctions Prédites
SEQ ID n° 4482	2346	142	SA-1000.1	202238-203305 m	similar to glutamyl-aminopeptidase and/or to endo-1,4-beta-glucanase
SEQ ID n° 4483	2347	143	SA-1001.1	203490-203729 m	Unknown
SEQ ID n° 4484	2348	144	SA-1002.1	203890-204174 p	similar to unknown protein
SEQ ID n° 4485	2349	145	SA-1003.1	204171-204494 p	similar to thioredoxin H1
SEQ ID n° 4486	2350	146	SA-1004.1	204527-205153 p	weakly similar to phenylalanyl-tRNA synthetase (beta subunit)
SEQ ID n° 4487	2351	147	SA-1006.1	205207-205923 m	similar to unknown proteins
SEQ ID n° 4488	2352	148	SA-1007.1	206004-206399 p	similar to single-strand binding protein
SEQ ID n° 4489	2353	149	SA-101.1	2190539-2191378 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4490	2354	150	SA-1010.1	206522-207166 p	similar to unknown proteins
SEQ ID n° 4491	2355	151	SA-1012.1	207193-208938 p	Similar to two-component sensor histidine kinase
SEQ ID n° 4492	2356	152	SA-1013.1	208919-209659 p	Similar to two-component response regulator
SEQ ID n° 4493	2357	154	SA-1016.1	209829-210284 p	similar to unknown proteins
SEQ ID n° 4494	2358	155	SA-1017.1	210286-211014 p	similar to unknown proteins
SEQ ID n° 4495	2359	156	SA-1018.1	211257-212885 p	Similar to (oligopeptide) ABC transporter (binding protein)
SEQ ID n° 4496	2360	157	SA-1019.1	212998-213975 p	Similar to oligopeptide ABC transporter (permease)
SEQ ID n° 4497	2361	158	SA-102.1	2191378-2191917 m	similar to phosphatidylglycerophosphate synthase
SEQ ID n° 4498	2362	159	SA-1020.1	213972-214793 p	Similar to ABC transporter (permease)
SEQ ID n° 4499	2363	160	SA-1021.1	214805-215608 p	Similar to oligopeptide ABC transporter (ATP-binding protein)
SEQ ID n° 4500	2364	161	SA-1022.2	215592-216218 p	Similar to oligopeptide ABC transporter (ATP-binding protein)
SEQ ID n° 4501	2365	163	SA-1024.2	10743-10877 p	Hypothetical protein
SEQ ID n° 4502	2366	164	SA-1025.2	10369-10740 p	similar to unknown protein
SEQ ID n° 4503	2367	165	SA-1026.1	10110-10382 p	similar to unknown protein
SEQ ID n° 4504	2368	166	SA-1028.1	6322-9819 p	similar to transcription-repair coupling factor
SEQ ID n° 4505	2369	167	SA-1029.1	5750-6325 p	similar to peptidyl-tRNA hydrolase
SEQ ID n° 4506	2370	168	SA-103.1	2192043-2193326 m	similar to unknown proteins
SEQ ID n° 4507	2371	169	SA-1030.1	4551-5666 p	similar to hypothetical GTP binding protein
SEQ ID n° 4508	2372	170	SA-1031.1	4188-4391 p	Hypothetical protein
SEQ ID n° 4509	2373	171	SA-1032.1	3789-3986 p	similar to unknown protein
SEQ ID n° 4510	2374	172	SA-1033.1	2898-3779 p	similar to putative transcription regulator
SEQ ID n° 4511	2375	173	SA-1034.1	1692-2828 p	beta subunit of DNA polymerase III

SEQ ID n° 4512	2376	174	SA-1035.2	176-1537 p	replication initiation protein DnaA
SEQ ID n° 4513	2377	175	SA-1036.2	2217128-2217901 p	similar to chromosome partitioning protein ParB
SEQ ID n° 4514	2378	176	SA-1037.1	2215801-2217030 p	similar to serine protease
SEQ ID n° 4515	2379	177	SA-1038.1	2215121-2215600 m	similar to unknown protein
SEQ ID n° 4516	2380	178	SA-1039.1	2211963-2214542 p	similar to unknown transmembrane protein
SEQ ID n° 4517	2381	179	SA-104.1	2193328-2194572 m	similar to unknown proteins
SEQ ID n° 4518	2382	180	SA-1041.1	2210221-2211840 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4519	2383	181	SA-1042.4	2209283-2210155 p	similar to unknown proteins
SEQ ID n° 4520	2384	182	SA-1044.2	891360-892031 p	similar to thiamine-phosphate pyrophosphorylase (ThiE)
SEQ ID n° 4521	2385	183	SA-1045.1	892158-893417 p	Similar to UDP-N-acetylglucosamine 1-carboxyvinyltransferase
SEQ ID n° 4522	2386	184	SA-1046.1	893519-894073 p	Similar to acetyl transferase
SEQ ID n° 4523	2387	185	SA-1047.1	894066-895349 p	Similar to unknown proteins
SEQ ID n° 4524	2388	186	SA-1048.1	895372-896232 p	similar to methionine aminopeptidase
SEQ ID n° 4525	2389	187	SA-1049.1	896234-897154 p	similar to unknown proteins
SEQ ID n° 4526	2390	188	SA-105.1	2194811-2195164 p	similar to unknown proteins
SEQ ID n° 4527	2391	189	SA-1050.1	897171-897626 m	similar to unknown proteins
SEQ ID n° 4528	2392	190	SA-1051.1	897808-898317 p	similar to unknown proteins
SEQ ID n° 4529	2393	191	SA-1053.1	898457-900415 p	similar to DNA ligase
SEQ ID n° 4530	2394	192	SA-1054.1	900427-901446 p	similar to unknown proteins
SEQ ID n° 4531	2395	193	SA-1055.1	901450-903750 p	Similar to pullulanase
SEQ ID n° 4532	2396	194	SA-1056.1	903956-905824 p	similar to 1,4-alpha-glucan branching enzyme
SEQ ID n° 4533	2397	195	SA-1057.1	905866-907005 p	similar to glucose-1-phosphate adenyltransferase
SEQ ID n° 4534	2398	196	SA-1058.1	906995-908128 p	similar to glucose-1-phosphate adenyltransferase
SEQ ID n° 4535	2399	197	SA-1059.2	908125-909555 p	similar to glycogen synthase
SEQ ID n° 4536	2400	198	SA-106.1	2195167-2196276 p	Similar to recombination protein RecF
SEQ ID n° 4537	2401	201	SA-1062.2	1167875-1169149 m	similar to xanthine permease
SEQ ID n° 4538	2402	202	SA-1063.1	1166988-1167776 p	similar to nitrite or formate transporter
SEQ ID n° 4539	2403	203	SA-1064.1	1165654-1166889 p	similar to unknown protein (N-terminal part)
SEQ ID n° 4540	2404	204	SA-1065.1	1165028-1165630 p	Unknown, similar to unknown protein
SEQ ID n° 4541	2405	205	SA-1066.1	1164072-1165010 p	similar to other protein (putative thiamine biosynthesis lipoprotein)
SEQ ID n° 4542	2406	206	SA-1067.1	1163741-1163923 p	similar to 4-oxalocrotonate tautomerase
SEQ ID n° 4543	2407	207	SA-1068.1	1163034-1163603 m	thymidine kinase
SEQ ID n° 4544	2408	208	SA-1069.1	1161920-1162999 m	peptide chain release factor 1
SEQ ID n° 4545	2409	209	SA-107.1	2196288-2197154 m	putative transmembrane protein similar to unknown protein

SEQ ID n° 4546	2410	210	SA-1071.1	1161090-1161920 m	similar to protoporphyrinogen oxidase
SEQ ID n° 4547	2411	211	SA-1072.1	1160501-1161097 m	similar to unknown protein
SEQ ID n° 4548	2412	212	SA-1073.1	1159153-1160409 m	serine hydroxymethyltransferase
SEQ ID n° 4549	2413	213	SA-1074.1	1158171-1159148 m	similar to unknown protein
SEQ ID n° 4550	2414	214	SA-1075.1	1157567-1158169 m	similar to unknown protein
SEQ ID n° 4551	2415	215	SA-1076.1	1155834-1157555 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4552	2416	216	SA-1077.1	1154100-1155833 m	similar to ABC-transporter (ATP-binding protein)
SEQ ID n° 4553	2417	217	SA-1078.2	1153889-1154086 m	Unknown
SEQ ID n° 4554	2418	218	SA-1079.1	1152107-1153825 p	similar to phosphoglucomutase
SEQ ID n° 4555	2419	219	SA-108.1	2197210-2197872 m	similar to transcriptional regulator
SEQ ID n° 4556	2420	220	SA-1080.1	1151425-1151997 p	similar to unknown protein
SEQ ID n° 4557	2421	221	SA-1081.1	1150837-1151379 p	similar to pantothenate metabolism flavoprotein
SEQ ID n° 4558	2422	222	SA-1082.1	1150158-1150844 p	similar to flavoprotein
SEQ ID n° 4559	2423	223	SA-1083.1	1149168-1150004 m	similar to unknown protein
SEQ ID n° 4560	2424	224	SA-1084.1	1147976-1149175 m	similar to NADH-dependent oxidoreductase
SEQ ID n° 4561	2425	225	SA-1085.1	1146993-1147979 m	similar to unknown protein
SEQ ID n° 4562	2426	226	SA-1086.1	1146632-1146964 m	similar to glycine cleavage system H protein
SEQ ID n° 4563	2427	227	SA-1087.2	1145821-1146639 m	similar to unknown protein
SEQ ID n° 4564	2428	228	SA-1088.2	346400-347701 p	Similar to cellobiose-specific PTS enzyme IIC
SEQ ID n° 4565	2429	229	SA-1089.1	346078-346398 p	Similar to cellobiose-specific PTS enzyme IIB
SEQ ID n° 4566	2430	230	SA-109.1	2197947-2199428 m	similar to inosine monophosphate dehydrogenase
SEQ ID n° 4567	2431	231	SA-1090.1	345738-346061 p	Similar to cellobiose-specific PTS enzyme IIA
SEQ ID n° 4568	2432	232	SA-1091.2	344551-345534 p	similar to hypothetical transcriptional regulator
SEQ ID n° 4569	2433	233	SA-1092.2	343677-344432 p	similar to transcriptional regulator (DeoR family)
SEQ ID n° 4570	2434	234	SA-1093.1	342780-343556 m	similar to pyruvate formate-lyase activating enzyme
SEQ ID n° 4571	2435	235	SA-1094.1	342377-342733 p	similar to general stress protein GSP13
SEQ ID n° 4572	2436	236	SA-1095.1	340975-342375 p	C-terminal part similar to peptidyl-prolyl cis-trans isomerase
SEQ ID n° 4573	2437	237	SA-1096.1	340292-340933 p	Similar to two-component response regulator
SEQ ID n° 4574	2438	238	SA-1097.1	339280-340299 p	Similar to two-component sensor histidine kinase
SEQ ID n° 4575	2439	239	SA-1098.1	338568-339283 p	similar to unknown proteins
SEQ ID n° 4576	2440	241	SA-11.1	1031773-1034922 m	similar to plasmid proteins
SEQ ID n° 4577	2441	242	SA-110.1	2199585-2200070 m	similar to arginine repressor (ArgR)
SEQ ID n° 4578	2442	243	SA-1100.1	336474-338429 p	similar to hypothetical protein kinases
SEQ ID n° 4579	2443	244	SA-1101.1	335737-336474 p	similar to other proteins
SEQ ID n° 4580	2444	245	SA-1102.1	334377-335699 p	similar to RNA-binding protein Sun
SEQ ID n° 4581	2445	246	SA-1103.1	333452-334387 p	methionyl tRNA formyltransferase

SEQ ID n° 4582	2446	247	SA-1104.1	331015-333405 p	primosomal replication factor Y (primosomal protein N)
SEQ ID n° 4583	2447	248	SA-1105.2	330630-330941 p	similar to unknown proteins
SEQ ID n° 4584	2448	249	SA-1107.2	141808-143196 p	similar to argininosuccinate lyase
SEQ ID n° 4585	2449	250	SA-1108.1	140599-141789 p	similar to argininosuccinate synthase
SEQ ID n° 4586	2450	251	SA-1109.1	139375-140445 p	similar to two-component sensor histidine kinase
SEQ ID n° 4587	2451	252	SA-111.1	220080-2200760 m	similar to transcriptional regulator (Crp/Fnr family)
SEQ ID n° 4588	2452	253	SA-1110.1	138711-139382 p	similar to similar to two-component response regulator
SEQ ID n° 4589	2453	254	SA-1111.1	137994-138680 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4590	2454	255	SA-1112.1	136906-137994 p	similar to putative ABC transporter (permease)
SEQ ID n° 4591	2455	256	SA-1113.1	135738-136724 m	similar to transcription repressor of ribose operon
SEQ ID n° 4592	2456	257	SA-1114.1	134834-135745 m	similar to ribokinase
SEQ ID n° 4593	2457	258	SA-1115.1	134461-134859 m	similar to ribose ABC transporter (permease)
SEQ ID n° 4594	2458	259	SA-1116.1	132967-134445 m	similar to ribose ABC transporter (ATP-binding protein)
SEQ ID n° 4595	2459	260	SA-1118.1	132024-132965 m	similar to ribose ABC transporter (permease)
SEQ ID n° 4596	2460	261	SA-1119.1	131003-131971 m	similar to ribose ABC transporter (binding protein)
SEQ ID n° 4597	2461	262	SA-112.1	220068-2201672 p	similar to unknown proteins
SEQ ID n° 4598	2462	263	SA-1120.1	129509-130963 p	glutamyl-tRNA synthetase
SEQ ID n° 4599	2463	264	SA-1121.1	127981-129300 p	similar to mercury(II) reductase
SEQ ID n° 4600	2464	265	SA-1122.1	127330-127827 p	similar to carbonic anhydrase
SEQ ID n° 4601	2465	266	SA-1123.2	125830-127194 p	similar to DNA repair protein RadA
SEQ ID n° 4602	2466	268	SA-1125.2	1963678-1964112 m	similar to N-acetylgalactosamine-specific PTS system enzyme
SEQ ID n° 4603	2467	269	SA-1126.2	1962479-1963675 m	IIA similar to glucuronyl hydrolase
SEQ ID n° 4604	2468	270	SA-1127.1	1961930-1962424 m	similar to N-acetylgalactosamine-specific PTS system enzyme
SEQ ID n° 4605	2469	271	SA-1129.1	1961028-1961894 m	IIA similar to N-acetylgalactosamine-specific PTS system enzyme
SEQ ID n° 4606	2470	272	SA-113.1	2201939-2203171 p	IIC similar to arginine deiminase
SEQ ID n° 4607	2471	273	SA-1131.1	1960226-1961041 m	similar to N-acetylgalactosamine-specific PTS system enzyme
SEQ ID n° 4608	2472	274	SA-1132.1	1958242-1960146 m	IID similar to heparitin-sulfate lyase from Flavobacterium heparinum
SEQ ID n° 4609	2473	275	SA-1134.1	1957161-1958162 m	similar to transcriptional regulator (LacI family)
SEQ ID n° 4610	2474	276	SA-1135.1	1956476-1957090 p	similar to polypeptide deformylase
SEQ ID n° 4611	2475	277	SA-1136.1	1955766-1956410 p	similar to transcriptional regulator (Crp/Fnr family)
SEQ ID n° 4612	2476	278	SA-1137.1	1954491-1955654 m	similar to permease (major facilitator superfamily)

SEQ ID n° 4613	2477	279	SA-1138.1	1953462-1954445 p	similar to oxidoreductase
SEQ ID n° 4614	2478	280	SA-1139.1	1951419-1953314 m	similar to endopeptidase
SEQ ID n° 4615	2479	281	SA-114.1	2203267-2203608 p	similar to unknown proteins
SEQ ID n° 4616	2480	282	SA-1140.1	1950416-1951369 p	similar to L,D-carboxypeptidase LdcA
SEQ ID n° 4617	2481	283	SA-1141.1	1950013-1950249 m	lysine rich protein
SEQ ID n° 4618	2482	284	SA-1143.2	1947942-1950011 m	similar to Na ⁺ /H ⁺ antiporter
SEQ ID n° 4619	2483	285	SA-1144.1	1946133-1947665 p	similar to alkyl hydroperoxide reductase (large subunit) and NADH dehydrogenase
SEQ ID n° 4620	2484	286	SA-1145.2	1945555-1946115 p	similar to alkyl hydroperoxide reductase (small subunit)
SEQ ID n° 4621	2485	288	SA-1148.2	1764745-1765479 m	similar to unknown proteins
SEQ ID n° 4622	2486	289	SA-1149.1	1764471-1764722 m	Unknown
SEQ ID n° 4623	2487	290	SA-115.1	2203624-2204637 p	similar to ornithine transcarbamylase
SEQ ID n° 4624	2488	291	SA-1150.1	1763365-1764474 m	Similar to unknown proteins
SEQ ID n° 4625	2489	292	SA-1151.1	1762559-1763272 m	Similar to transcriptional regulator
SEQ ID n° 4626	2490	293	SA-1152.1	1762016-1762420 m	Similar to unknown proteins
SEQ ID n° 4627	2491	294	SA-1153.1	1761037-1761750 p	Similar to glycerol uptake facilitator
SEQ ID n° 4628	2492	295	SA-1154.1	1760553-1761027 p	Similar to unknown proteins
SEQ ID n° 4629	2493	296	SA-1155.1	1760075-1760653 p	Similar to unknown proteins
SEQ ID n° 4630	2494	297	SA-1156.1	1759048-1760037 p	Similar to hypothetical dihydroxyacetone kinase
SEQ ID n° 4631	2495	298	SA-1157.1	1758338-1758874 m	Similar to putative transcription regulator
SEQ ID n° 4632	2496	299	SA-1158.1	1757342-1758328 m	Similar to putative dihydroxyacetone kinase
SEQ ID n° 4633	2497	300	SA-116.1	2204700-2206127 p	similar to arginine/ornithine antiporter
SEQ ID n° 4634	2498	301	SA-1160.1	1756508-1757224 m	Similar to unknown proteins
SEQ ID n° 4635	2499	302	SA-1161.1	1755664-1756353 m	Similar to other proteins
SEQ ID n° 4636	2500	303	SA-1162.1	1754729-1755562 m	similar to (amino acid ?) ABC transporter (binding protein)
SEQ ID n° 4637	2501	304	SA-1163.1	1753764-1754588 m	Similar to unknown protein
SEQ ID n° 4638	2502	305	SA-1165.1	1752254-1753630 m	Similar to unknown proteins
SEQ ID n° 4639	2503	307	SA-1167.1	1751191-1752261 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4640	2504	308	SA-1169.1	1750502-1751194 m	similar to ABC transporter (permease)
SEQ ID n° 4641	2505	309	SA-117.1	2206148-2207104 p	similar to carbamate kinase
SEQ ID n° 4642	2506	310	SA-1170.1	1749398-1750453 p	similar to NAD alcohol dehydrogenase
SEQ ID n° 4643	2507	311	SA-1171.2	1747917-1749284 p	Similar to branched-chain amino acid transporter
SEQ ID n° 4644	2508	312	SA-1173.3	1370419-1370727 m	truncated C-terminal part
SEQ ID n° 4645	2509	313	SA-1174.1	1370724-1371032 m	Unknown
SEQ ID n° 4646	2510	314	SA-1175.1	1371034-1371606 m	truncated C-terminal part
SEQ ID n° 4647	2511	315	SA-1176.1	1371607-1372110 m	truncated N-terminal part

SEQ ID n° 4648	2512	316	SA-1177.1	1372110-1372424 m	Unknown
SEQ ID n° 4649	2513	317	SA-1178.1	1372659-1373864 m	similar to unknown proteins
SEQ ID n° 4650	2514	318	SA-1179.1	1373916-1374407 m	Unknown
SEQ ID n° 4651	2515	319	SA-118.2	2207214-2208239 m	tryptophanyl-tRNA synthetase
SEQ ID n° 4652	2516	320	SA-1180.2	1374421-1378032 m	similar to probable type II DNA modification enzyme
SEQ ID n° 4653	2517	321	SA-1182.1	1378369-1379568 m	similar to integrase
SEQ ID n° 4654	2518	322	SA-1183.1	1379571-1379888 m	similar to unknown proteins
SEQ ID n° 4655	2519	323	SA-1184.1	1380538-1381431 m	similar to unknown proteins
SEQ ID n° 4656	2520	324	SA-1187.1	1381531-1382937 m	similar to Phospho-beta-D-galactosidase
SEQ ID n° 4657	2521	326	SA-1189.1	1383024-1384730 m	Similar to lactose-specific PTS system enzyme IIBC
SEQ ID n° 4658	2522	328	SA-1190.1	1384730-1385047 m	Similar to lactose-specific PTS system enzyme IIA
SEQ ID n° 4659	2523	329	SA-1191.1	1385076-1385909 m	Similar to transcription antiterminator
SEQ ID n° 4660	2524	330	SA-1192.1	1386302-1387282 m	similar to tagatose-1,6-diphosphate aldolase
SEQ ID n° 4661	2525	331	SA-1193.1	1387287-1388216 m	similar to tagatose 6-phosphate kinase
SEQ ID n° 4662	2526	332	SA-1194.3	1388229-1388744 m	similar to galactose-6-phosphate isomerase (large subunit)
SEQ ID n° 4663	2527	333	SA-1195.3	1388761-1389186 m	similar to galactose 6-P isomerase (small subunit)
SEQ ID n° 4664	2528	334	SA-1196.2	248617-249024 p	Unknown
SEQ ID n° 4665	2529	335	SA-1197.1	248225-248515 p	similar to unknown proteins
SEQ ID n° 4666	2530	336	SA-1198.1	247430-247717 m	similar to unknown proteins
SEQ ID n° 4667	2531	337	SA-1199.1	247105-247440 m	similar to unknown protein
SEQ ID n° 4668	2532	340	SA-1201.1	246330-246800 p	Unknown
SEQ ID n° 4669	2533	342	SA-1203.1	244906-246162 p	similar to plasmid recombination / mobilization protein
SEQ ID n° 4670	2534	343	SA-1205.1	244155-244589 p	Unknown
SEQ ID n° 4671	2535	344	SA-1207.1	243263-244051 p	similar to plasmid replication protein
SEQ ID n° 4672	2536	345	SA-1208.1	242322-242963 p	Unknown
SEQ ID n° 4673	2537	346	SA-1209.1	241992-242318 p	similar to replication protein (N-terminal part)
SEQ ID n° 4674	2538	348	SA-1210.1	241705-241986 p	Unknown
SEQ ID n° 4675	2539	349	SA-1211.1	241271-241564 p	Unknown
SEQ ID n° 4676	2540	350	SA-1212.1	240555-241106 m	similar to transcriptional regulator
SEQ ID n° 4677	2541	351	SA-1213.1	239332-240486 m	similar to integrase
SEQ ID n° 4678	2542	352	SA-1214.2	238812-239204 p	ribosomal protein S9
SEQ ID n° 4679	2543	353	SA-1215.2	238345-238791 p	50S ribosomal protein L13
SEQ ID n° 4680	2544	354	SA-1216.2	951207-951863 p	similar to unknown proteins
SEQ ID n° 4681	2545	355	SA-1217.1	950447-951217 p	similar to N-acetylglucosamine metabolism protein
SEQ ID n° 4682	2546	356	SA-1218.1	949709-950446 p	similar to other proteins
SEQ ID n° 4683	2547	357	SA-1219.1	948575-949705 p	similar to coproporphyrinogen III oxidase (HemN)

SEQ ID n° 4684	2548	358	SA-1220.1	948103-948483 p	similar to unknown proteins
SEQ ID n° 4685	2549	359	SA-1221.1	947707-948078 p	similar to unknown proteins
SEQ ID n° 4686	2550	360	SA-1222.1	946232-947584 p	similar to phospho-sugar mutase
SEQ ID n° 4687	2551	361	SA-1223.1	945219-946178 p	similar to unknown proteins
SEQ ID n° 4688	2552	362	SA-1225.1	944365-945216 p	similar to unknown proteins
SEQ ID n° 4689	2553	363	SA-1226.1	942882-944225 m	similar to unknown protein (to hypothetical UDP-N-acetylmuramyl tripeptide synthetase??) -----
SEQ ID n° 4690	2554	364	SA-1227.1	942097-942882 m	similar to other proteins
SEQ ID n° 4691	2555	365	SA-1228.1	941001-941990 p	similar to lipoate-protein ligase A
SEQ ID n° 4692	2556	366	SA-123.1	90455-90841 p	50S ribosomal protein L17
SEQ ID n° 4693	2557	367	SA-1230.1	939146-940903 p	similar to acetoin dehydrogenase E3 component (dihydrolipoamide dehydrogenase)
SEQ ID n° 4694	2558	368	SA-1231.1	937698-939086 p	similar to acetoin dehydrogenase E2 component (dihydrolipoamide acetyltransferase)
SEQ ID n° 4695	2559	369	SA-1232.1	936573-937571 p	similar to acetoin dehydrogenase (TPP-dependent) beta chain
SEQ ID n° 4696	2560	371	SA-1234.1	935530-936498 p	similar to acetoin dehydrogenase (TPP-dependent) alpha chain
SEQ ID n° 4697	2561	372	SA-1236.3	933471-935381 p	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4698	2562	373	SA-1238.2	1190234-1190593 m	Unknown
SEQ ID n° 4699	2563	374	SA-1239.1	1190655-1191113 m	Unknown
SEQ ID n° 4700	2564	375	SA-1240.1	1191174-1191401 m	Unknown
SEQ ID n° 4701	2565	378	SA-1243.1	1191645-1192142 m	Unknown
SEQ ID n° 4702	2566	379	SA-1244.1	1192156-1192908 m	Unknown
SEQ ID n° 4703	2567	380	SA-1245.1	1192957-1193442 m	Unknown
SEQ ID n° 4704	2568	381	SA-1246.1	1193548-1193775 m	Unknown
SEQ ID n° 4705	2569	382	SA-1247.1	1194050-1196848 m	putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 4706	2570	383	SA-1248.1	1196915-1197625 m	putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 4707	2571	385	SA-125.1	89502-90440 p	RNA polymerase (alpha subunit)
SEQ ID n° 4708	2572	386	SA-1250.1	1197640-1199871 m	similar to surface exclusion protein Sec10 precursor - Enterococcus faecalis plasmid pCF10
SEQ ID n° 4709	2573	387	SA-1251.1	1199888-1200187 m	Unknown
SEQ ID n° 4710	2574	388	SA-1252.1	1200570-1200767 m	Similar to phage protein
SEQ ID n° 4711	2575	389	SA-1253.1	1200764-1200949 m	Unknown
SEQ ID n° 4712	2576	390	SA-1254.1	1200951-1201976 m	similar to putative plasmid replication initiator protein A
SEQ ID n° 4713	2577	391	SA-1255.1	1201978-1202142 m	Unknown

SEQ ID n° 4714	2578	392	SA-1256.1	1202249-1202692 m	Unknown
SEQ ID n° 4715	2579	393	SA-1257.1	1202958-1203215 m	Similar to unknown protein
SEQ ID n° 4716	2580	394	SA-1258.1	1203399-1203770 m	Unknown
SEQ ID n° 4717	2581	395	SA-1261.1	1203884-1204867 m	similar to GMP reductase
SEQ ID n° 4718	2582	396	SA-1265.2	1001539-1002666 m	similar to Na ⁺ /H ⁺ antiporter
SEQ ID n° 4719	2583	397	SA-1266.1	1000513-1001475 p	similar to unknown protein
SEQ ID n° 4720	2584	398	SA-1267.1	999946-1000359 p	similar to unknown protein
SEQ ID n° 4721	2585	399	SA-1268.1	999187-999930 p	similar to sortase
SEQ ID n° 4722	2586	400	SA-127.2	89069-89452 p	30S ribosomal protein S11
SEQ ID n° 4723	2587	401	SA-1271.1	996721-999180 p	DNA gyrase A subunit
SEQ ID n° 4724	2588	402	SA-1273.1	995493-996482 m	similar to L-Lactate Dehydrogenase
SEQ ID n° 4725	2589	403	SA-1274.1	994003-995373 p	similar to NADH Oxidase
SEQ ID n° 4726	2590	404	SA-1275.1	992841-993797 p	similar to sugar (?) ABC transporter (permease)
SEQ ID n° 4727	2591	405	SA-1278.1	991778-992839 p	similar to sugar (?) ABC transporter (permease)
SEQ ID n° 4728	2592	406	SA-1279.1	990250-991785 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4729	2593	407	SA-128.3	88686-89051 p	ribosomal protein S13
SEQ ID n° 4730	2594	408	SA-1280.1	989056-990105 p	similar to ABC transporter (binding protein)
SEQ ID n° 4731	2595	409	SA-1281.1	988602-988991 p	similar to cytidine deaminase
SEQ ID n° 4732	2596	410	SA-1282.1	987686-988276 p	similar to E. coli 16S RNA m2G1207 methyltransferase
SEQ ID n° 4733	2597	411	SA-1283.1	986660-987580 m	similar to pantothenate kinase
SEQ ID n° 4734	2598	412	SA-1284.1	986358-986591 m	30S ribosomal protein S20
SEQ ID n° 4735	2599	413	SA-1285.1	985443-986273 p	similar to amino acid ABC transporter (binding protein)
SEQ ID n° 4736	2600	414	SA-1286.3	984797-985426 p	similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 4737	2601	416	SA-1288.3	682243-683688 m	similar to unknown proteins
SEQ ID n° 4738	2602	417	SA-1289.1	681515-682186 p	similar to ABC transporter (permease)
SEQ ID n° 4739	2603	418	SA-1290.1	681109-681303 p	Unknown
SEQ ID n° 4740	2604	419	SA-1291.1	680078-681100 p	Similar to ABC transporter (permease) (N-terminal part)
SEQ ID n° 4741	2605	420	SA-1292.1	679319-680068 p	Similar to ABC transporter (permease) (C-terminal part)
SEQ ID n° 4742	2606	421	SA-1293.1	678697-679272 p	Unknown
SEQ ID n° 4743	2607	422	SA-1295.1	677517-678728 p	similar to glycosyltransferases
SEQ ID n° 4744	2608	423	SA-1296.1	675317-677512 p	similar to acyl-carrier-protein synthase
SEQ ID n° 4745	2609	424	SA-1297.1	674367-675320 p	similar to aminomethyltransferase
SEQ ID n° 4746	2610	425	SA-1298.1	672367-674370 p	Unknown
SEQ ID n° 4747	2611	426	SA-1299.1	671492-672370 p	similar to ABC transporter (permease)
SEQ ID n° 4748	2612	427	SA-13.1	1034915-1035172 m	Unknown
SEQ ID n° 4749	2613	429	SA-1300.1	670570-671499 p	similar to ABC transporter (ATP-binding protein)

SEQ ID n° 4750	2614	430	SA-1301.1	670104-670580 p	similar to (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
SEQ ID n° 4751	2615	431	SA-1302.1	669815-670120 p	similar to acyl carrier protein
SEQ ID n° 4752	2616	432	SA-1303.1	669100-669822 p	similar to 3-oxoacyl-(acyl-carrier protein) reductase
SEQ ID n° 4753	2617	433	SA-1304.2	668255-669103 p	weakly similar to malonyl-CoA:Acyl carrier protein transacylase
SEQ ID n° 4754	2618	434	SA-1305.2	667950-668255 p	Unknown
SEQ ID n° 4755	2619	435	SA-1308.2	1846687-1847346 p	Similar to transcriptional regulator (FNR/CRP family)
SEQ ID n° 4756	2620	436	SA-1309.1	1847370-1849655 p	Similar to X-prolyl dipeptidyl aminopeptidase
SEQ ID n° 4757	2621	437	SA-1310.1	1849659-1850018 p	Similar to unknown proteins
SEQ ID n° 4758	2622	438	SA-1311.1	1850064-1851044 p	Similar to heptaprenyl diphosphate synthase component II
SEQ ID n° 4759	2623	439	SA-1312.1	1851129-1852877 m	Similar to cytochrome D ABC transporter ATP binding and permease protein
SEQ ID n° 4760	2624	440	SA-1313.1	1852870-1854588 m	Similar to cytochrome D ABC transporter (ATP binding and permease protein)
SEQ ID n° 4761	2625	441	SA-1314.1	1854588-1855607 m	Similar to cytochrome bd ubiquinol oxidase (subunit II)
SEQ ID n° 4762	2626	442	SA-1316.1	1855608-1857035 m	Similar to cytochrome bd ubiquinol oxidase (subunit I)
SEQ ID n° 4763	2627	443	SA-1318.1	1857138-1858346 m	Similar to hypothetical NADH dehydrogenase
SEQ ID n° 4764	2628	444	SA-1319.2	1858359-1859258 m	Similar to hypothetical menaquinone biosynthesis protein MenA
SEQ ID n° 4765	2629	446	SA-1320.1	1108449-1108991 m	Unknown
SEQ ID n° 4766	2630	447	SA-1322.1	1108996-1109481 m	Unknown
SEQ ID n° 4767	2631	449	SA-1324.1	1109840-1110130 m	similar to immunogenic secreted protein precursor (C-terminal part)
SEQ ID n° 4768	2632	450	SA-1325.1	1110093-1110989 m	similar to immunogenic secreted protein precursor (internal part)
SEQ ID n° 4769	2633	451	SA-1326.1	1110908-1111336 m	Unknown
SEQ ID n° 4770	2634	452	SA-1327.1	1111463-1111717 m	similar to unknown protein
SEQ ID n° 4771	2635	453	SA-1328.1	1111738-1112328 m	similar to other protein
SEQ ID n° 4772	2636	454	SA-1329.1	1112342-1112647 m	Unknown
SEQ ID n° 4773	2637	456	SA-1330.1	1112776-1113690 m	similar to unknown protein
SEQ ID n° 4774	2638	457	SA-1331.1	1113693-1114049 m	Unknown
SEQ ID n° 4775	2639	458	SA-1332.1	1114061-1114411 m	similar to unknown protein
SEQ ID n° 4776	2640	460	SA-1334.1	1114425-1118396 m	weakly similar to DNA-translocase
SEQ ID n° 4777	2641	461	SA-1335.1	1118329-1118832 m	Unknown
SEQ ID n° 4778	2642	462	SA-1336.1	1118839-1120113 m	similar to unknown protein

SEQ ID n° 4779	2643	463	SA-1337.1	1120113-1120355 m	similar to unknown protein
SEQ ID n° 4780	2644	464	SA-1338.1	1120339-1120812 m	Unknown
SEQ ID n° 4781	2645	465	SA-1339.3	1120821-1123932 m	similar to unknown protein
SEQ ID n° 4782	2646	467	SA-1340.2	1964390-1965202 p	similar to 5-keto-D-gluconate 5-reductase
SEQ ID n° 4783	2647	468	SA-1341.1	1965219-1965857 p	similar to sugar-phosphate isomerase
SEQ ID n° 4784	2648	469	SA-1342.1	1965883-1966890 p	similar to 2-keto-3-deoxygluconate kinase
SEQ ID n° 4785	2649	470	SA-1343.1	1966902-1967540 p	similar to 2-dehydro-3-deoxyphosphogluconate aldolase
SEQ ID n° 4786	2650	471	SA-1344.2	1968248-1970101 p	Unknown
SEQ ID n° 4787	2651	472	SA-1345.1	1970272-1970886 m	similar to unknown protein
SEQ ID n° 4788	2652	473	SA-1347.1	1971019-1971444 m	similar to transcription regulator (MarR family)
SEQ ID n° 4789	2653	474	SA-1348.1	1971571-1975977 m	DNA polymerase III (alpha subunit)
SEQ ID n° 4790	2654	475	SA-1349.1	1976100-1976684 m	similar to peptidoglycan hydrolase
SEQ ID n° 4791	2655	477	SA-1350.1	1976809-1978662 m	prolyl-tRNA synthetase
SEQ ID n° 4792	2656	478	SA-1351.1	1978754-1980013 m	similar to other protein
SEQ ID n° 4793	2657	479	SA-1352.1	1980044-1980838 m	similar to phosphatidate cytidyltransferase (CDP-diglyceride synthase)
SEQ ID n° 4794	2658	480	SA-1353.3	1980853-1981605 m	similar to undecaprenyl pyrophosphate synthetase
SEQ ID n° 4795	2659	481	SA-1354.2	1218666-1219799 m	similar to hypothetical permeases
SEQ ID n° 4796	2660	482	SA-1355.1	1219890-1221326 m	similar to 6-phospho-beta-glucosidase
SEQ ID n° 4797	2661	483	SA-1356.1	1221489-1222103 m	similar to other proteins
SEQ ID n° 4798	2662	484	SA-1357.1	1222204-1223025 m	similar to unknown protein
SEQ ID n° 4799	2663	485	SA-1358.1	1223650-1224579 p	similar to transcriptional regulator (AraC/XylS family)
SEQ ID n° 4800	2664	486	SA-1359.1	1224636-1226168 m	similar to putative channel transporter
SEQ ID n° 4801	2665	487	SA-136.1	264314-265096 p	Unknown
SEQ ID n° 4802	2666	488	SA-1361.3	1226277-1227350 m	similar to spermidine/putrescine ABC transporter (binding protein)
SEQ ID n° 4803	2667	489	SA-1362.1	1227343-1228119 m	similar to spermidine/putrescine ABC transporter (permease)
SEQ ID n° 4804	2668	490	SA-1363.1	1228116-1228910 m	similar to spermidine/putrescine ABC transporter (permease)
SEQ ID n° 4805	2669	491	SA-1364.1	1228894-1230048 m	similar to spermidine/putrescine ABC transporter (ATP-binding protein)
SEQ ID n° 4806	2670	492	SA-1365.1	1230097-1230999 m	similar to UDP-N-acetylenolpyruvoylglucosamine reductase
SEQ ID n° 4807	2671	493	SA-1366.1	1231143-1231631 m	similar to hydroxymethylpterin pyrophosphokinase
SEQ ID n° 4808	2672	494	SA-1367.1	1231628-1231990 m	similar to dihydroneopterin aldolase
SEQ ID n° 4809	2673	495	SA-1368.1	1231992-1232795 m	similar to dihydropteroate synthase
SEQ ID n° 4810	2674	496	SA-137.1	265093-266226 m	similar to integrase
SEQ ID n° 4811	2675	497	SA-1370.1	1232799-1233362 m	similar to GTP cyclohydrolase

SEQ ID n° 4812	2676	498	SA-1371.1	1233381-1234643 m	similar to folyl-polyglutamate synthetase
SEQ ID n° 4813	2677	499	SA-1372.1	1234645-1235532 m	similar to unknown protein
SEQ ID n° 4814	2678	500	SA-1373.2	1235519-1236385 m	similar to homoserine kinase
SEQ ID n° 4815	2679	501	SA-1374.3	1236387-1237670 m	similar to homoserine dehydrogenase
SEQ ID n° 4816	2680	502	SA-1377.1	1647764-1649380 m	Similar to nickel ABC transporter (binding protein)
SEQ ID n° 4817	2681	503	SA-1378.1	1646833-1647777 m	Similar to nickel ABC transporter (permease)
SEQ ID n° 4818	2682	504	SA-1379.1	1646027-1646833 m	Similar to oligopeptide and nickel ABC transporter (permease)
SEQ ID n° 4819	2683	505	SA-138.1	266230-266436 m	similar to unknown proteins
SEQ ID n° 4820	2684	506	SA-1380.1	1645251-1646039 m	Similar to oligopeptide ABC transporter (ATP-binding protein)
SEQ ID n° 4821	2685	507	SA-1381.1	1644584-1645264 m	Similar to oligopeptide or nickel ABC transporter (ATP-binding protein)
SEQ ID n° 4822	2686	508	SA-1382.1	1643735-1644463 m	uridylate kinase
SEQ ID n° 4823	2687	509	SA-1383.1	1643162-1643719 m	ribosome recycling factor
SEQ ID n° 4824	2688	510	SA-1384.1	1642190-1643044 m	similar to unknown proteins
SEQ ID n° 4825	2689	511	SA-1385.1	1641543-1642052 m	similar to peptide methionine sulfoxide reductase
SEQ ID n° 4826	2690	512	SA-1386.1	1641331-1641546 m	similar to unknown proteins
SEQ ID n° 4827	2691	513	SA-1388.1	1639401-1641173 m	similar to myosin-crossreactive streptococcal antigen (unknown function)
SEQ ID n° 4828	2692	514	SA-1389.1	1638291-1639277 m	similar to phosphate starvation inducible protein, unknown function
SEQ ID n° 4829	2693	515	SA-1390.2	1637397-1638200 p	Unknown
SEQ ID n° 4830	2694	516	SA-1391.2	2028876-2030531 m	similar to two-component sensor histidine kinase
SEQ ID n° 4831	2695	517	SA-1392.1	2030524-2031201 m	similar to two-component response regulator (PhoB)
SEQ ID n° 4832	2696	518	SA-1393.1	2031201-2031857 m	Similar to transcriptional regulator PhoU
SEQ ID n° 4833	2697	519	SA-1394.1	2031854-2032603 m	similar to phosphate ABC transporter (ATP-binding protein)
SEQ ID n° 4834	2698	520	SA-1395.1	2032596-2033474 m	similar to phosphate ABC transporter (permease)
SEQ ID n° 4835	2699	521	SA-1396.1	2033476-2034321 m	similar to phosphate ABC transporter (permease)
SEQ ID n° 4836	2700	523	SA-1399.1	2034336-2035217 m	similar to phosphate ABC transporter (binding protein)
SEQ ID n° 4837	2701	524	SA-14.1	1035400-1035831 m	similar to single-strand binding protein
SEQ ID n° 4838	2702	525	SA-140.1	266442-267230 m	similar to replication protein (plasmid)
SEQ ID n° 4839	2703	526	SA-1400.1	2035418-2036005 m	Unknown
SEQ ID n° 4840	2704	527	SA-1401.1	2036002-2036742 m	Similar to unknown proteins
SEQ ID n° 4841	2705	528	SA-1403.1	2036742-2037695 m	Similar to ribosomal protein L11 methyltransferase
SEQ ID n° 4842	2706	529	SA-1404.1	2037692-2038150 m	Similar to unknown proteins
SEQ ID n° 4843	2707	530	SA-1405.1	2038274-2038990 p	Similar to transcriptional regulators

SEQ ID n° 4844	2708	531	SA-1406.1	2039029-2039499 m	Similar to unknown proteins
SEQ ID n° 4845	2709	532	SA-1407.1	2039471-2039929 m	Similar to unknown proteins
SEQ ID n° 4846	2710	533	SA-1408.1	2039916-2040059 m	Unknown
SEQ ID n° 4847	2711	534	SA-141.1	267411-267725 m	Unknown
SEQ ID n° 4848	2712	535	SA-1410.1	2040065-2040535 m	Similar to unknown proteins
SEQ ID n° 4849	2713	536	SA-1413.1	1907193-1909571 m	similar to xylulose-5-phosphate/fructose-6-phosphate phosphoketolase
SEQ ID n° 4850	2714	537	SA-1414.1	1909661-1910752 m	Similar to unknown proteins
SEQ ID n° 4851	2715	538	SA-1415.1	1911069-1912748 m	Weakly similar to transcriptional regulator (antiterminator)
SEQ ID n° 4852	2716	539	SA-1416.2	1912814-1913575 m	similar to unknown proteins
SEQ ID n° 4853	2717	541	SA-1418.2	1913595-1915112 m	similar to L-xylulose kinase
SEQ ID n° 4854	2718	542	SA-1419.1	1915178-1916167 m	Unknown
SEQ ID n° 4855	2719	543	SA-1421.1	1916189-1917640 m	Similar to galactitol-specific PTS enzyme IIC
SEQ ID n° 4856	2720	544	SA-1424.1	1917738-1918694 m	similar to glycerate dehydrogenase
SEQ ID n° 4857	2721	545	SA-1425.2	1918713-1919732 m	Unknown
SEQ ID n° 4858	2722	546	SA-1427.1	154278-155117 p	similar to undecaprenol kinase
SEQ ID n° 4859	2723	547	SA-1429.1	155237-155992 p	similar to competence negative regulator mecA
SEQ ID n° 4860	2724	548	SA-143.1	267824-269182 m	weakly similar to DNA translocase
SEQ ID n° 4861	2725	549	SA-1430.1	155994-157154 p	similar to Streptococcus mutans RgpG protein required for biosynthesis of rhamnose-glucose polysaccharide
SEQ ID n° 4862	2726	550	SA-1431.1	157319-158089 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4863	2727	551	SA-1432.1	158126-159388 p	similar to unknown protein
SEQ ID n° 4864	2728	552	SA-1433.1	159390-160622 p	similar to aminotransferase, putative cysteine desulfurase
SEQ ID n° 4865	2729	553	SA-1434.1	160609-161052 p	similar to NifU protein
SEQ ID n° 4866	2730	554	SA-1435.1	161152-162570 p	similar to unknown protein
SEQ ID n° 4867	2731	555	SA-1436.1	162642-163829 m	similar to serine-type D-Ala-D-Ala carboxypeptidase (penicillin binding protein)
SEQ ID n° 4868	2732	556	SA-1437.2	163982-165217 m	similar to serine-type D-Ala-D-Ala carboxypeptidase (penicillin binding protein)
SEQ ID n° 4869	2733	557	SA-1438.2	613213-613473 m	50S ribosomal protein L31
SEQ ID n° 4870	2734	558	SA-1439.1	611580-613100 p	similar to metal ABC transporter (binding protein)
SEQ ID n° 4871	2735	559	SA-144.1	269145-269570 m	Unknown
SEQ ID n° 4872	2736	560	SA-1440.1	610041-611438 p	similar to dipeptidase
SEQ ID n° 4873	2737	561	SA-1441.1	609115-610026 p	Similar to unknown proteins
SEQ ID n° 4874	2738	562	SA-1442.2	608144-609118 p	Similar to unknown proteins
SEQ ID n° 4875	2739	563	SA-1445.2	607257-608147 p	Similar to unknown proteins

SEQ ID n° 4876	2740	564	SA-1446.1	606690-607103 m	Similar to unknown proteins
SEQ ID n° 4877	2741	565	SA-1447.1	605461-606444 m	Similar to purine nucleoside hydrolase
SEQ ID n° 4878	2742	566	SA-1448.1	604907-605464 m	Similar to unknown proteins
SEQ ID n° 4879	2743	567	SA-1449.1	603521-604867 p	asparaginyl-tRNA synthetase
SEQ ID n° 4880	2744	568	SA-145.1	269636-269917 m	similar to unknown proteins
SEQ ID n° 4881	2745	569	SA-1450.1	602307-603500 p	similar to aspartate aminotransferase
SEQ ID n° 4882	2746	570	SA-1451.1	599714-602221 p	similar to ATP-dependent DNA helicase
SEQ ID n° 4883	2747	571	SA-1453.2	598842-599606 p	Similar to meso-2,3-butanediol dehydrogenase (D-acetoin forming)
SEQ ID n° 4884	2748	573	SA-1457.2	1931763-1934015 p	similar to gamma-glutamylcysteine synthetase (for the N-terminal part) and to cyanophycinsynthetase (C-terminal part)
SEQ ID n° 4885	2749	574	SA-1458.1	1930919-1931593 p	Similar to unknown proteins
SEQ ID n° 4886	2750	575	SA-146.1	270084-270542 p	Unknown
SEQ ID n° 4887	2751	576	SA-1460.1	1929535-1930557 m	Similar to transcriptional regulators and to PTS enzyme II
SEQ ID n° 4888	2752	577	SA-1462.1	1927872-1929164 m	Similar to adenylosuccinate synthase
SEQ ID n° 4889	2753	578	SA-1463.1	1927556-1927849 p	Unknown
SEQ ID n° 4890	2754	579	SA-1465.1	1926600-1927247 m	Similar to unknown protein
SEQ ID n° 4891	2755	580	SA-1466.1	1924993-1926432 m	Putative PTS enzyme IIC
SEQ ID n° 4892	2756	581	SA-1467.1	1924687-1924965 m	similar to mannitol-specific PTS enzyme IIB
SEQ ID n° 4893	2757	582	SA-1468.1	1924135-1924620 m	similar to mannitol-specific PTS enzyme IIA
SEQ ID n° 4894	2758	583	SA-1469.1	1923357-1924022 m	Similar to hexulose-6-phosphate synthase
SEQ ID n° 4895	2759	584	SA-147.1	270919-271479 p	similar to putative acetyl transferase
SEQ ID n° 4896	2760	585	SA-1470.1	1922490-1923353 m	Similar to hexulose-6-phosphate isomerase
SEQ ID n° 4897	2761	586	SA-1471.1	1921772-1922488 m	Similar to L-ribulose-5-phosphate 4-epimerase
SEQ ID n° 4898	2762	587	SA-1472.3	1921114-1921761 m	similar to hypothetical transaldolase
SEQ ID n° 4899	2763	589	SA-1474.1	104233-104421 p	Unknown
SEQ ID n° 4900	2764	590	SA-1475.1	104755-105075 p	Unknown
SEQ ID n° 4901	2765	591	SA-1477.1	105865-106416 p	Similar to unknown proteins
SEQ ID n° 4902	2766	592	SA-1478.1	106636-107055 p	Similar to unknown proteins
SEQ ID n° 4903	2767	593	SA-1479.1	107255-107734 p	Similar to the putative sigma factor ComX1
SEQ ID n° 4904	2768	594	SA-148.1	271579-272157 p	similar to putative acetyl transferase
SEQ ID n° 4905	2769	595	SA-1480.1	107856-108548 p	Similar to phosphoglycerate mutase
SEQ ID n° 4906	2770	596	SA-1481.1	108545-109297 p	Similar to D,D-carboxypeptidase
SEQ ID n° 4907	2771	597	SA-1482.1	109294-109869 p	Similar to N-acetyl muramidase

SEQ ID n° 4908	2772	598	SA-1483.1	110012-111046 p	Similar to transcription repressor of class I heat-shock (HrcA)
SEQ ID n° 4909	2773	599	SA-1484.1	111049-111621 p	Similar to heat shock protein GrpE
SEQ ID n° 4910	2774	600	SA-1486.3	111802-113631 p	Chaperone protein DnaK
SEQ ID n° 4911	2775	601	SA-1487.3	113920-115059 p	Chaperone protein DnaJ
SEQ ID n° 4912	2776	604	SA-149.1	272173-272880 p	similar to unknown protein
SEQ ID n° 4913	2777	605	SA-1490.2	1650563-1650988 m	ribosomal protein L11
SEQ ID n° 4914	2778	606	SA-1491.1	1651199-1652581 m	similar to transmembrane efflux pump protein
SEQ ID n° 4915	2779	607	SA-1492.1	1652589-1653803 m	similar to hypothetical N-acyl-L-amino acid amidohydrolase
SEQ ID n° 4916	2780	608	SA-1493.1	1654062-1654967 p	similar to transcriptional regulator (LysR family)
SEQ ID n° 4917	2781	609	SA-1494.1	1655028-1655381 p	similar to unknown proteins
SEQ ID n° 4918	2782	610	SA-1495.1	1655410-1657125 m	similar to para-aminobenzoate synthase, component I
SEQ ID n° 4919	2783	611	SA-1498.1	1657207-1659648 m	Similar to DNA translocase
SEQ ID n° 4920	2784	612	SA-1499.1	1659823-1660626 p	similar to peptidyl-prolyl cis-trans isomerase
					similar to surface antigen proteins, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 4921	2785	613	SA-15.3	745991-748792 p	
SEQ ID n° 4922	2786	614	SA-150.1	273140-274087 m	similar to unknown proteins
SEQ ID n° 4923	2787	615	SA-1500.1	1660678-1661511 m	similar to metal ABC transporter (permease)
SEQ ID n° 4924	2788	616	SA-1501.1	1661513-1662229 m	similar to metal ABC transporter (ATP-binding protein)
SEQ ID n° 4925	2789	617	SA-1502.2	1662400-1663326 m	similar to metal ABC transporter (binding protein)
					similar to C5A peptidase, putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 4926	2790	618	SA-1503.2	2080002-2084714 p	
SEQ ID n° 4927	2791	619	SA-1504.1	2084947-2085633 p	similar to two-component response regulator
SEQ ID n° 4928	2792	620	SA-1506.1	2085630-2087018 p	similar to two-component sensor histidine kinase
SEQ ID n° 4929	2793	621	SA-1507.1	2087111-2087719 p	Unknown
SEQ ID n° 4930	2794	622	SA-151.1	274084-274575 m	similar to RNA polymerase ECF sigma factor
SEQ ID n° 4931	2795	623	SA-1511.2	2087806-2090307 m	leucyl-tRNA synthetase
SEQ ID n° 4932	2796	624	SA-1512.2	356072-356626 p	similar to unknown proteins
SEQ ID n° 4933	2797	625	SA-1513.1	355330-355995 p	similar to competence protein ComFC
SEQ ID n° 4934	2798	627	SA-1515.1	354041-355330 p	competence protein ComFA
SEQ ID n° 4935	2799	628	SA-1516.1	353341-353985 m	similar to unknown proteins
SEQ ID n° 4936	2800	629	SA-1517.1	352324-353250 m	similar to cysteine synthetase A
SEQ ID n° 4937	2801	630	SA-1518.1	351084-352172 p	similar to glycerol dehydrogenase
SEQ ID n° 4938	2802	631	SA-1519.1	350348-351016 p	similar to hypothetical transaldolase
SEQ ID n° 4939	2803	632	SA-152.1	275157-275765 m	similar to transcriptional regulator (TetR/AcrR family)
SEQ ID n° 4940	2804	633	SA-1520.3	347882-350338 p	similar to formate acetyltransferase

SEQ ID n° 4941	2805	634	SA-1521.1	1269394-1269624 m	similar to unknown proteins
SEQ ID n° 4942	2806	635	SA-1522.1	1269707-1270729 m	similar to branched-chain-amino-acid aminotransferase
SEQ ID n° 4943	2807	636	SA-1523.1	1270842-1273301 m	Similar to topoisomerase IV subunit A
SEQ ID n° 4944	2808	637	SA-1524.1	1273435-1275384 m	similar to DNA topoisomerase IV (subunit B)
SEQ ID n° 4945	2809	638	SA-1525.1	1275498-1276136 p	similar to unknown proteins
SEQ ID n° 4946	2810	639	SA-1526.1	1276202-1276855 m	Uracil DNA glycosylase
SEQ ID n° 4947	2811	640	SA-1527.1	1276954-1277439 m	similar to unknown proteins
SEQ ID n° 4948	2812	641	SA-1528.1	1277553-1278794 m	CMP-N-acetylneuraminic acid synthetase
SEQ ID n° 4949	2813	642	SA-153.1	275789-276886 m	Similar to ABC transporter (permease)
SEQ ID n° 4950	2814	643	SA-1530.1	1278805-1279434 m	similar to glycosylation or acetylation protein
SEQ ID n° 4951	2815	644	SA-1531.1	1279431-1280585 m	similar to UDP-N-acetylglucosamine-2-epimerase
SEQ ID n° 4952	2816	645	SA-1532.1	1280662-1281687 m	similar to N-acetylneuraminic acid synthetase
SEQ ID n° 4953	2817	646	SA-1533.1	1281687-1283087 m	capsular polysaccharide repeat unit transporter
SEQ ID n° 4954	2818	647	SA-1535.2	973749-974447 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4955	2819	648	SA-1536.1	972955-973737 m	putative ABC transporter (permease)
SEQ ID n° 4956	2820	649	SA-1537.1	972314-972919 p	similar to unknown transmembrane protein
SEQ ID n° 4957	2821	650	SA-1538.1	970960-971529 m	similar to acetyltransferase (chloramphenicol ?)
SEQ ID n° 4958	2822	651	SA-1539.1	969736-970839 p	Unknown
SEQ ID n° 4959	2823	653	SA-1543.1	966944-969736 p	similar to cation-transporting P-ATPase
SEQ ID n° 4960	2824	654	SA-1544.1	966358-966792 p	similar to peptide methionine sulfoxide reductase
SEQ ID n° 4961	2825	655	SA-1545.1	965894-966358 p	similar to unknown proteins
SEQ ID n° 4962	2826	656	SA-1546.1	965232-965885 p	similar to unknown proteins
SEQ ID n° 4963	2827	657	SA-1547.2	962493-965126 p	weakly similar to histidine triad protein, putative lipoprotein
SEQ ID n° 4964	2828	658	SA-155.1	276889-277605 m	1.2 Transport/binding proteins and lipoproteins
SEQ ID n° 4965	2829	659	SA-1551.2	656055-658727 p	Putative peptidoglycan bound protein (FPXTG motif)
SEQ ID n° 4966	2830	660	SA-1552.1	658766-659008 p	similar to hypothetical sortase protein (N-terminal part)
SEQ ID n° 4967	2831	661	SA-1554.1	658966-659571 p	similar to putative surface protein (sortase)
SEQ ID n° 4968	2832	662	SA-1555.1	659787-660425 p	Unknown
SEQ ID n° 4969	2833	664	SA-1558.1	661523-661699 p	Unknown
SEQ ID n° 4970	2834	665	SA-156.2	277869-278552 m	similar to unknown proteins
SEQ ID n° 4971	2835	666	SA-1560.1	663083-663313 p	Unknown
SEQ ID n° 4972	2836	667	SA-1562.1	663374-664525 p	Putative cell wall protein, weakly similar to peptidase or esterase
SEQ ID n° 4973	2837	668	SA-1563.1	664725-665717 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4974	2838	669	SA-1564.1	665720-666538 p	similar to unknown proteins, putative transmembrane protein
SEQ ID n° 4975	2839	670	SA-1565.2	666540-667325 p	similar to unknown proteins, putative transmembrane protein

SEQ ID n° 4976	2840	671	SA-1567.3	1831820-1832131 m	Similar to unknown proteins
SEQ ID n° 4977	2841	672	SA-1568.3	1832242-1833135 p	Similar to Ribonuclease HII
SEQ ID n° 4978	2842	673	SA-157.2	278573-278884 m	similar to unknown proteins
SEQ ID n° 4979	2843	674	SA-1570.1	1833151-1833744 p	Similar to signal peptidase 1
SEQ ID n° 4980	2844	675	SA-1571.2	1833873-1836293 p	Similar to putative exodeoxyribonuclease V
SEQ ID n° 4981	2845	676	SA-1573.3	1836407-1836889 p	Similar to unknown proteins
SEQ ID n° 4982	2846	677	SA-1574.1	1836960-1838054 m	Similar to DNA-damage-inducible protein P
SEQ ID n° 4983	2847	678	SA-1577.1	1838239-1840551 p	Similar to Pyruvate formate-lyase
SEQ ID n° 4984	2848	679	SA-1578.1	1840652-1841026 p	Similar to FMN-binding protein
SEQ ID n° 4985	2849	680	SA-1579.1	1841061-1841990 m	Similar to unknown proteins
SEQ ID n° 4986	2850	681	SA-1580.1	1841987-1842742 m	Similar to C3-degrading proteinase
SEQ ID n° 4987	2851	682	SA-1582.1	1842867-1843763 p	Similar to unknown protein
SEQ ID n° 4988	2852	683	SA-1584.3	1843842-1844690 m	Similar to glycerol uptake facilitator
SEQ ID n° 4989	2853	684	SA-1585.3	1844859-1845311 m	Similar to unknown protein
SEQ ID n° 4990	2854	685	SA-1586.2	1845329-1846531 m	Similar to efflux protein
SEQ ID n° 4991	2855	686	SA-1587.2	624205-625581 p	similar to ABC transporter (permease)
SEQ ID n° 4992	2856	687	SA-1588.1	623549-624205 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 4993	2857	688	SA-1589.1	622262-623539 p	similar to ABC transporter (permease)
SEQ ID n° 4994	2858	689	SA-159.2	279076-279783 p	similar to other proteins
SEQ ID n° 4995	2859	690	SA-1590.1	621359-621520 m	similar to unknown proteins
SEQ ID n° 4996	2860	691	SA-1591.1	620914-621186 p	similar to transposase
SEQ ID n° 4997	2861	692	SA-1593.1	620591-620851 p	similar to transposase
SEQ ID n° 4998	2862	693	SA-1594.1	620369-620569 p	similar to transposase
SEQ ID n° 4999	2863	694	SA-1595.1	619803-620207 p	similar to insertion element protein
SEQ ID n° 5000	2864	695	SA-1596.1	618920-619540 m	similar to integrase C-terminal part (truncated)
SEQ ID n° 5001	2865	696	SA-1597.1	618428-618775 p	ribosomal protein L19
SEQ ID n° 5002	2866	697	SA-1598.1	616652-617848 p	similar to unknown proteins
SEQ ID n° 5003	2867	698	SA-1599.1	616384-616659 p	similar to other proteins
SEQ ID n° 5004	2868	700	SA-160.2	280061-281209 p	similar to N-acetylglucosamine-6-phosphate deacetylase
SEQ ID n° 5005	2869	701	SA-1601.1	615864-616307 p	similar to flavodoxin
SEQ ID n° 5006	2870	702	SA-1602.1	614783-615805 p	similar to adenosine deaminase
SEQ ID n° 5007	2871	703	SA-1603.2	613582-614517 m	similar to unknown proteins
SEQ ID n° 5008	2872	704	SA-1604.2	1867026-1867295 p	30S ribosomal protein S14
SEQ ID n° 5009	2873	705	SA-1605.2	1867657-1868682 m	similar to low specificity L-threonine aldolase
SEQ ID n° 5010	2874	706	SA-1606.2	1868802-1869812 m	Similar to other proteins (includin putative glycoprotein endopeptidase)

SEQ ID n° 5011	2875	707	SA-1607.1	1869888-1870334 m	Similar to putative acetyltransferase
SEQ ID n° 5012	2876	708	SA-1608.1	1870297-1870986 m	Similar to other proteins (includin putative glycoprotein endopeptidase)
SEQ ID n° 5013	2877	709	SA-1609.1	1871168-1871398 p	Similar to unknown proteins
SEQ ID n° 5014	2878	710	SA-161.2	1457718-1458158 m	Similar to hypothetical transcriptional regulator
SEQ ID n° 5015	2879	711	SA-1610.1	1871452-1873131 p	Similar to unknown proteins
SEQ ID n° 5016	2880	712	SA-1611.1	1873293-1873802 m	Similar to secreted unknown proteins -
SEQ ID n° 5017	2881	713	SA-1612.1	1873950-1875296 m	glutamine synthetase
SEQ ID n° 5018	2882	714	SA-1613.1	1875330-1875701 m	Similar to transcriptional regulator, MerR/GlnR family
SEQ ID n° 5019	2883	715	SA-1614.1	1875781-1876320 m	Similar to unknown protein
SEQ ID n° 5020	2884	717	SA-1616.1	1876583-1877779 m	Similar to phosphoglycerate kinase
SEQ ID n° 5021	2885	718	SA-1617.2	1877914-1878783 m	Similar to unknown lipoprotein
SEQ ID n° 5022	2886	719	SA-1619.2	2071134-2071673 p	similar to other proteins
SEQ ID n° 5023	2887	720	SA-162.1	1455854-1457671 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5024	2888	721	SA-1620.1	2071874-2072959 p	similar to glycerol dehydrogenase
SEQ ID n° 5025	2889	722	SA-1621.1	2073124-2073831 m	similar to unknown protein
SEQ ID n° 5026	2890	723	SA-1622.1	2074178-2076022 m	similar to other proteins
SEQ ID n° 5027	2891	724	SA-1623.1	2076067-2078304 m	similar to cobalamin-independent methionine synthase MetC
SEQ ID n° 5028	2892	725	SA-1624.1	2078674-2078997 m	similar to unknown protein
SEQ ID n° 5029	2893	726	SA-1625.2	2078987-2079679 m	similar to putative transport protein
SEQ ID n° 5030	2894	727	SA-1626.3	517632-518201 p	Similar to unknown proteins
SEQ ID n° 5031	2895	728	SA-1627.1	518299-518883 p	Similar to unknown proteins
SEQ ID n° 5032	2896	729	SA-1628.1	518880-519446 p	Similar to unknown proteins
SEQ ID n° 5033	2897	730	SA-1629.1	519446-522100 p	valyl-tRNA synthetase
SEQ ID n° 5034	2898	731	SA-163.1	1454110-1455864 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5035	2899	732	SA-1630.1	522336-523265 m	Similar to unknown proteins
SEQ ID n° 5036	2900	733	SA-1632.1	523682-524641 p	Similar to oxidoreductase
SEQ ID n° 5037	2901	734	SA-1634.1	524802-525704 p	Similar to putative divalent cation transport protein
SEQ ID n° 5038	2902	735	SA-1635.3	525864-526928 p	similar to unknown proteins
SEQ ID n° 5039	2903	736	SA-1636.2	1746531-1747739 m	Similar to transport proteins
SEQ ID n° 5040	2904	737	SA-1638.1	1745814-1746362 m	Similar to unknown proteins
SEQ ID n° 5041	2905	738	SA-164.1	1453376-1454002 m	similar to other proteins
SEQ ID n° 5042	2906	739	SA-1640.2	1744121-1745797 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5043	2907	740	SA-1641.2	1743298-1744128 m	Similar to hypothetical ABC transporter (permease)
SEQ ID n° 5044	2908	741	SA-1642.1	1742592-1743263 p	Similar to unknown proteins
SEQ ID n° 5045	2909	742	SA-1643.1	1741200-1742579 p	similar to Na ⁺ -transporting ATP synthase

SEQ ID n° 5046	2910	743	SA-1644.1	1740472-1741185 p	Similar to glucose-inhibited division protein GidB
SEQ ID n° 5047	2911	744	SA-1645.1	1739790-1740344 m	Similar to unknown proteins
SEQ ID n° 5048	2912	745	SA-1646.1	1738879-1739769 m	Similar to other proteins
SEQ ID n° 5049	2913	746	SA-1647.1	1738241-1738774 m	Similar to unknown proteins
SEQ ID n° 5050	2914	747	SA-1648.1	1737317-1738006 m	Similar to two-component response regulator
SEQ ID n° 5051	2915	748	SA-1649.2	1735822-1737327 m	Similar to two-component sensor histidine kinase
SEQ ID n° 5052	2916	749	SA-165.1	1452688-1453374 m	Similar to two-component response regulator
SEQ ID n° 5053	2917	750	SA-1650.2	1636729-1637205 m	similar to other proteins
SEQ ID n° 5054	2918	751	SA-1651.1	1635601-1636407 m	similar to unknown proteins
SEQ ID n° 5055	2919	752	SA-1653.1	1634709-1635206 m	similar to unknown proteins
SEQ ID n° 5056	2920	753	SA-1654.1	1634330-1634728 m	similar to diacylglycerol kinase
SEQ ID n° 5057	2921	754	SA-1655.1	1633389-1634288 m	similar to GTP binding proteins
SEQ ID n° 5058	2922	755	SA-1656.1	1632751-1633152 m	Unknown
SEQ ID n° 5059	2923	757	SA-1658.1	1631767-1632021 m	Unknown
SEQ ID n° 5060	2924	759	SA-166.1	1451459-1452688 m	Similar to two-component sensor histidine kinase
SEQ ID n° 5061	2925	760	SA-1660.1	1629232-1630824 p	Unknown
SEQ ID n° 5062	2926	761	SA-1662.1	1628167-1629015 m	similar to transcriptional regulator
SEQ ID n° 5063	2927	762	SA-1663.1	1627197-1628018 m	similar to formamidopyrimidine-DNA glycosylase
SEQ ID n° 5064	2928	763	SA-1665.1	1626613-1627200 m	similar to dephosphoerythrose A kinase
SEQ ID n° 5065	2929	764	SA-1667.1	1625788-1626489 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5066	2930	765	SA-1669.1	1624152-1625798 m	Unknown
SEQ ID n° 5067	2931	766	SA-167.1	1450465-1451343 m	similar to mevalonate kinase
SEQ ID n° 5068	2932	767	SA-1671.2	1622787-1623995 m	similar to efflux pump
SEQ ID n° 5069	2933	768	SA-1673.2	1988716-1990323 m	Similar to glucan 1,6-alpha-glucosidase
SEQ ID n° 5070	2934	769	SA-1674.1	1987633-1988628 m	Similar to UDP-galactose 4-epimerase
SEQ ID n° 5071	2935	770	SA-1675.1	1986932-1987612 p	similar to two-component response regulator
SEQ ID n° 5072	2936	771	SA-1676.1	1985404-1986930 p	similar to two-component sensor histidine kinase
SEQ ID n° 5073	2937	772	SA-1677.1	1983913-1985250 m	similar to organic acid transport protein
SEQ ID n° 5074	2938	774	SA-1679.4	1982725-1983888 m	similar to malic enzyme ((S)-malate:NAD+ oxidoreductase)
SEQ ID n° 5075	2939	775	SA-168.1	1449539-1450483 m	similar to mevalonate diphosphate decarboxylase
SEQ ID n° 5076	2940	776	SA-1680.2	1942117-1942581 m	Similar to transcriptional regulator CtsR
SEQ ID n° 5077	2941	777	SA-1681.2	1939673-1942120 m	Similar to endopeptidase Clp ATP-binding chain C
SEQ ID n° 5078	2942	778	SA-1682.1	1938962-1939453 p	Similar to acyltransferase
SEQ ID n° 5079	2943	779	SA-1683.1	1938307-1938948 p	similar to deoxyuridine kinase subunit
SEQ ID n° 5080	2944	780	SA-1685.1	1937209-1938186 p	Similar to transcription regulator
SEQ ID n° 5081	2945	781	SA-1686.1	1936350-1937225 p	Similar to other proteins

SEQ ID n° 5082	2946	782	SA-1687.1	1934954-1936210 p	Similar to other proteins
SEQ ID n° 5083	2947	783	SA-1689.3	1934139-1934957 p	Similar to unknown proteins
SEQ ID n° 5084	2948	784	SA-169.1	1448554-1449546 m	similar to phosphomevalonate kinase
SEQ ID n° 5085	2949	786	SA-1694.1	28975-30318 p	cell wall separation
SEQ ID n° 5086	2950	788	SA-1696.1	31518-32693 p	similar to aminotransferase
SEQ ID n° 5087	2951	789	SA-1697.2	32683-33444 p	similar to unknown protein
SEQ ID n° 5088	2952	790	SA-1698.2	1725796-1726299 m	Similar to unknown proteins
SEQ ID n° 5089	2953	791	SA-1699.1	1726385-1727716 m	Similar to UDP-N-acetylmuramate-alanine ligase
SEQ ID n° 5090	2954	792	SA-17.1	1038856-1039578 m	similar to unknown protein, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5091	2955	793	SA-170.1	1447562-1448557 m	similar to Isopentenyl diphosphate isomerase
SEQ ID n° 5092	2956	794	SA-1700.1	1727726-1728343 m	Similar to unknown proteins
SEQ ID n° 5093	2957	795	SA-1702.1	1728508-1731606 m	Similar to SWI/SNF family helicase
SEQ ID n° 5094	2958	796	SA-1704.1	1731762-1733072 m	Similar to GTP binding protein
SEQ ID n° 5095	2959	797	SA-1705.1	1733120-1734022 m	Similar to primosome component (helicase loader) Dnal
SEQ ID n° 5096	2960	798	SA-1706.2	1734019-1735194 m	Similar to unknown proteins
SEQ ID n° 5097	2961	799	SA-1707.2	1735194-1735673 m	Similar to unknown protein
SEQ ID n° 5098	2962	800	SA-171.1	1447276-1447494 m	Unknown
SEQ ID n° 5099	2963	801	SA-1710.2	2102209-2103849 m	similar to unknown protein
SEQ ID n° 5100	2964	802	SA-1711.1	2101372-2102184 p	similar to other protein
SEQ ID n° 5101	2965	803	SA-1712.1	2098935-2101256 m	similar to penicillin-binding protein 2a
SEQ ID n° 5102	2966	804	SA-1714.1	2098525-2098698 m	similar to other protein
SEQ ID n° 5103	2967	806	SA-1716.2	2098282-2098213 m	putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5104	2968	807	SA-1718.2	1809587-1810966 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5105	2969	808	SA-1719.1	1810959-1811672 m	Similar to unknown protein, putative ABC transporter (permease)
SEQ ID n° 5106	2970	809	SA-172.1	1446376-1447230 p	similar to unknown proteins
SEQ ID n° 5107	2971	810	SA-1720.1	1811672-1812268 m	Similar to unknown protein
SEQ ID n° 5108	2972	811	SA-1722.1	1812279-1812710 m	Similar to unknown protein
SEQ ID n° 5109	2973	812	SA-1723.1	1812725-1812880 m	Unknown
SEQ ID n° 5110	2974	813	SA-1724.1	1812894-1813433 m	Similar to 2,3-dihydroxybenzoate-AMP ligase (C-terminal domain)
SEQ ID n° 5111	2975	814	SA-1725.1	1813437-1814261 m	Similar to 2,3-dihydroxybenzoate-AMP ligase (enterobactin synthetase component E)
SEQ ID n° 5112	2976	815	SA-1726.1	1814469-1814843 m	Similar to transcription regulator

SEQ ID n° 5113	2977	816	SA-1728.1	1814891-1815343 m	Similar to late competence protein required for DNA binding and uptake comEB
SEQ ID n° 5114	2978	817	SA-1729.1	1815355-1816422 m	Similar to similar to Xaa-Pro dipeptidase
SEQ ID n° 5115	2979	818	SA-173.1	144361-1446275 p	Similar to unknown proteins
SEQ ID n° 5116	2980	819	SA-1730.1	1816522-1817238 m	Similar to unknown protein
SEQ ID n° 5117	2981	820	SA-1731.2	1817240-1818739 m	Similar to multidrug resistance protein
SEQ ID n° 5118	2982	821	SA-1732.2	959863-960279 p	nucleoside-diphosphate kinase
SEQ ID n° 5119	2983	822	SA-1734.1	959293-959553 p	hypothetical
SEQ ID n° 5120	2984	824	SA-1737.1	957581-958246 p	similar to unknown proteins
SEQ ID n° 5121	2985	825	SA-1738.1	957253-957594 p	similar to unknown proteins
SEQ ID n° 5122	2986	826	SA-1739.1	956387-957256 p	similar to unknown proteins
SEQ ID n° 5123	2987	827	SA-174.1	1444618-1445274 p	similar to other proteins
SEQ ID n° 5124	2988	828	SA-1740.2	952252-956385 p	similar to unknown proteins
SEQ ID n° 5125	2989	829	SA-1741.2	2040862-2041353 p	Similar to putative acetyl transferase
SEQ ID n° 5126	2990	830	SA-1742.1	2041346-2042614 p	Similar to unknown proteins
SEQ ID n° 5127	2991	831	SA-1744.1	2043157-2043462 m	Unknown
SEQ ID n° 5128	2992	832	SA-1745.1	2043446-2043847 m	Unknown
SEQ ID n° 5129	2993	833	SA-1746.1	2043835-2045043 m	Similar to unknown proteins
SEQ ID n° 5130	2994	834	SA-1747.1	2045045-2045437 m	Unknown
SEQ ID n° 5131	2995	835	SA-1748.1	2045504-2045959 m	weakly similar to integrase
SEQ ID n° 5132	2996	837	SA-175.1	1444176-1444625 p	similar to unknown proteins
SEQ ID n° 5133	2997	838	SA-1750.1	2046386-2046622 m	hypothetical gene
SEQ ID n° 5134	2998	839	SA-1751.1	2046794-2047288 p	Similar to transcriptional regulator (phage related)
SEQ ID n° 5135	2999	840	SA-1753.2	774016-774840 p	Similar to unknown proteins
SEQ ID n° 5136	3000	841	SA-1754.1	773219-774016 p	Similar to unknown proteins
SEQ ID n° 5137	3001	842	SA-1755.1	769589-773128 p	Similar to chromosome segregation SMC protein
SEQ ID n° 5138	3002	843	SA-1756.1	768895-769581 p	Similar to ribonuclease III
SEQ ID n° 5139	3003	844	SA-1757.1	768351-768719 p	Similar to unknown proteins
SEQ ID n° 5140	3004	845	SA-1758.1	767539-768348 p	Similar to unknown proteins
SEQ ID n° 5141	3005	846	SA-1759.1	766186-767535 p	Similar to two-component sensor histidine kinase
SEQ ID n° 5142	3006	847	SA-176.1	1442801-1444084 p	similar to 3-hydroxy-3-methylglutaryl-coenzyme A reductase
SEQ ID n° 5143	3007	848	SA-1760.1	765483-766193 p	similar to two-component response regulator
SEQ ID n° 5144	3008	850	SA-1762.2	1503364-1503699 m	Similar to unknown proteins
SEQ ID n° 5145	3009	851	SA-1763.1	1503783-1504949 m	Similar to chorismate synthase
SEQ ID n° 5146	3010	852	SA-1764.1	1504950-1506017 m	Similar to 3-dehydroquinase synthase
SEQ ID n° 5147	3011	853	SA-1765.1	1506111-1506788 m	similar to 3-dehydroquinase dehydratase

SEQ ID n° 5148	3012	854	SA-1766.1	1506788-1507945 m	Similar to unknown proteins
SEQ ID n° 5149	3013	855	SA-1767.1	1508080-1510224 p	Similar to unknown proteins
SEQ ID n° 5150	3014	856	SA-1768.2	1510465-1510824 m	50S ribosomal protein L20
SEQ ID n° 5151	3015	857	SA-1769.2	1513077-1513274 p	Similar to ferredoxin
SEQ ID n° 5152	3016	858	SA-177.1	1441627-1442799 p	similar to 3-hydroxy-3-methylglutaryl-coenzyme A synthase (HMG-CoA synthase)
SEQ ID n° 5153	3017	859	SA-1770.1	1513261-1513752 m	Similar to unknown protein
SEQ ID n° 5154	3018	860	SA-1771.1	1513781-1515001 m	Similar to tripeptidase
SEQ ID n° 5155	3019	861	SA-1772.1	1515137-1516771 m	Similar to unknown protein
SEQ ID n° 5156	3020	862	SA-1773.1	1516895-1518349 p	Similar to UDP-N-acetyluramoylalanine-D-glutamyl-2, 6-diaminopimelate ligase
SEQ ID n° 5157	3021	863	SA-1774.2	1518504-1519298 p	similar to ferrichrome ABC transporter (ATP-binding protein)
SEQ ID n° 5158	3022	864	SA-1775.2	1519322-1520254 p	similar to ferrichrome ABC transporter (binding protein)
SEQ ID n° 5159	3023	865	SA-1776.1	1520270-1521295 p	similar to ferrichrome ABC transporter (permease)
SEQ ID n° 5160	3024	866	SA-1778.1	1521292-1522293 p	Similar to ferrichrome ABC transporter (permease)
SEQ ID n° 5161	3025	867	SA-1779.2	1522323-1522976 m	Similar to unknown protein
SEQ ID n° 5162	3026	868	SA-178.1	1440634-1441473 m	thymidylate synthase
SEQ ID n° 5163	3027	869	SA-1780.2	327377-328462 p	Similar to two-component sensor histidine kinase
SEQ ID n° 5164	3028	870	SA-1781.1	326640-327380 p	Similar to ABC transporter (permease)
SEQ ID n° 5165	3029	871	SA-1783.2	325757-326653 p	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5166	3030	872	SA-1784.2	324021-325628 p	similar to unknown proteins
SEQ ID n° 5167	3031	873	SA-1785.1	323350-323832 m	similar to autoinducer-2 production protein (LuxS)
SEQ ID n° 5168	3032	874	SA-1787.1	321662-323125 p	similar to unknown proteins
SEQ ID n° 5169	3033	875	SA-1788.1	320495-321649 p	similar to unknown proteins
SEQ ID n° 5170	3034	876	SA-179.1	1440060-1440554 m	similar to dihydrofolate reductase
SEQ ID n° 5171	3035	877	SA-1791.2	319693-320025 p	similar to unknown proteins
SEQ ID n° 5172	3036	878	SA-1792.2	1469012-1470220 m	similar to poly(A) polymerase
SEQ ID n° 5173	3037	879	SA-1793.1	1467133-1469001 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5174	3038	880	SA-1794.1	1466663-1467136 m	similar to unknown proteins
SEQ ID n° 5175	3039	881	SA-1795.1	1464857-1466596 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5176	3040	882	SA-1796.1	1463083-1464852 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5177	3041	883	SA-1797.2	1462533-1463042 p	similar to unknown proteins
SEQ ID n° 5178	3042	884	SA-1798.3	1568308-1569804 m	Similar to 4-alpha-glucanotransferase (amylomaltase)
SEQ ID n° 5179	3043	885	SA-1799.3	1566032-1568296 m	Similar to glycogen phosphorylase
SEQ ID n° 5180	3044	886	SA-18.1	1039655-1041916 m	similar to plasmid surface exclusion protein, putative peptidoglycan bound protein (LPXTG motif)

SEQ ID n° 5181	3045	887	SA-1801.2	572506-572721 p	similar to exodeoxyribonuclease VII small chain
SEQ ID n° 5182	3046	888	SA-1802.1	571188-572528 p	similar to exodeoxyribonuclease VII large chain
SEQ ID n° 5183	3047	889	SA-1803.1	570226-571062 p	similar to unknown proteins
SEQ ID n° 5184	3048	890	SA-1804.1	569375-570229 p	similar to bifunctional methylenetetrahydrofolate dehydrogenase / methylenetetrahydrofolate cyclohydrolase (Fold)
SEQ ID n° 5185	3049	891	SA-1806.1	567542-569236 p	similar to phosphomannomutase
SEQ ID n° 5186	3050	892	SA-1807.2	566521-567255 p	similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 5187	3051	893	SA-1808.3	565836-566528 p	Similar to amino acid ABC transporter (permease)
SEQ ID n° 5188	3052	894	SA-1809.2	1132887-1133588 m	similar to orotidine 5'-phosphate decarboxylase
SEQ ID n° 5189	3053	895	SA-181.1	1438616-1439842 m	similar to ATP-dependent Clp protease subunit X
SEQ ID n° 5190	3054	896	SA-1810.1	1132245-1132874 m	similar to orotate phosphoribosyltransferase
SEQ ID n° 5191	3055	897	SA-1812.1	1130941-1132233 m	similar to dihydroorotase
SEQ ID n° 5192	3056	898	SA-1813.1	1129853-1130776 m	similar to aspartate transcarbamoylase
SEQ ID n° 5193	3057	899	SA-1814.1	1128763-1129839 m	similar to glutaminase of carbamoyl-phosphate synthase
SEQ ID n° 5194	3058	900	SA-1815.1	1125550-1128732 m	similar to carbamoyl-phosphate synthase (glutamine-hydrolyzing)
SEQ ID n° 5195	3059	901	SA-1816.1	1125100-1125423 m	Unknown
SEQ ID n° 5196	3060	902	SA-1817.3	1124321-1125103 m	Unknown
SEQ ID n° 5197	3061	903	SA-1818.2	1714000-1714794 m	Similar to glutamate racemase
SEQ ID n° 5198	3062	904	SA-1819.1	1713029-1714003 m	Similar to unknown proteins
SEQ ID n° 5199	3063	905	SA-182.1	1438009-1438605 m	similar to hypothetical GTP-binding protein
SEQ ID n° 5200	3064	906	SA-1820.1	1712526-1713047 m	Similar to unknown proteins
SEQ ID n° 5201	3065	907	SA-1821.1	1712056-1712529 m	Similar to unknown protein
SEQ ID n° 5202	3066	908	SA-1822.1	1711326-1712066 m	Similar to integrase
SEQ ID n° 5203	3067	910	SA-1824.2	1710619-1711326 m	similar to unknown proteins
SEQ ID n° 5204	3068	911	SA-1825.2	1710038-1710622 m	similar to unknown proteins
SEQ ID n° 5205	3069	912	SA-1826.1	1709326-1710048 m	Similar to ribosomal large subunit pseudouridine synthase B (hypothetical)
SEQ ID n° 5206	3070	913	SA-1827.1	1709075-1709326 m	Similar to unknown protein
SEQ ID n° 5207	3071	915	SA-183.1	1437399-1438028 p	similar to transcriptional regulator
SEQ ID n° 5208	3072	916	SA-1830.1	1707612-1709051 p	Similar to TRK potassium uptake system protein TrkH
SEQ ID n° 5209	3073	917	SA-1832.2	1706258-1707607 p	Similar to TRK potassium uptake system protein TrkA
SEQ ID n° 5210	3074	918	SA-1834.1	2058938-2059792 m	similar to aminoglycoside 6-adenylyltransferase
SEQ ID n° 5211	3075	919	SA-1835.1	2059971-2060363 m	Unknown
SEQ ID n° 5212	3076	920	SA-1836.1	2060798-2061406 m	Unknown

SEQ ID n° 5213	3077	921	SA-1837.1	2061393-2061728 m	similar to unknown protein
SEQ ID n° 5214	3078	922	SA-1838.1	2062250-2062738 m	Unknown, similar to unknown protein
SEQ ID n° 5215	3079	923	SA-1839.1	2062907-2063668 m	similar to unknown protein (transmembrane)
SEQ ID n° 5216	3080	924	SA-1840.1	2063665-2064567 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5217	3081	925	SA-1841.1	2064772-2065848 m	similar to transcription regulator (N-terminal part)
SEQ ID n° 5218	3082	926	SA-1842.1	2066097-2066711 m	similar to unknown protein
SEQ ID n° 5219	3083	927	SA-1843.1	2066925-2067221 p	similar to unknown protein
SEQ ID n° 5220	3084	928	SA-1844.1	2067238-2067798 p	similar to unknown protein
SEQ ID n° 5221	3085	929	SA-1845.1	2068103-2068966 p	similar to unknown transmembrane protein
SEQ ID n° 5222	3086	930	SA-1846.1	2069123-2069425 m	similar to unknown protein
SEQ ID n° 5223	3087	931	SA-1847.2	2069598-2070365 m	CAMP factor
SEQ ID n° 5224	3088	932	SA-1849.1	1470405-1471253 m	similar to unknown proteins
SEQ ID n° 5225	3089	933	SA-185.1	1436804-1437127 m	Unknown
SEQ ID n° 5226	3090	934	SA-1850.1	1471253-1471633 m	similar to unknown proteins
SEQ ID n° 5227	3091	935	SA-1851.1	1471705-1472232 m	Unknown
SEQ ID n° 5228	3092	936	SA-1852.1	1472289-1472783 m	similar to unknown proteins
SEQ ID n° 5229	3093	938	SA-1855.1	1472972-1474936 m	Similar to fructose-specific PTS enzyme IIABC
SEQ ID n° 5230	3094	939	SA-1858.1	1474933-1475844 m	similar to Fructose-1-phosphate kinase
SEQ ID n° 5231	3095	940	SA-1859.1	1475841-1476650 m	similar to transcription repressor of fructose operon FruR
SEQ ID n° 5232	3096	941	SA-186.1	1435996-1436646 m	Unknown
SEQ ID n° 5233	3097	942	SA-1860.1	1476715-1477950 m	similar to Cell Wall Muropeptide Branching Enzyme
SEQ ID n° 5234	3098	943	SA-1861.1	1478094-1479725 m	similar to cell wall proteins, putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5235	3099	944	SA-1862.1	1479940-1480863 m	similar to 2-dehydropanoate 2-reductase
SEQ ID n° 5236	3100	946	SA-1866.2	833223-834167 p	Similar to transcriptional regulator (LacI family)
SEQ ID n° 5237	3101	947	SA-1867.1	834226-835218 m	similar to esterase
SEQ ID n° 5238	3102	948	SA-1868.1	835394-836122 p	Similar to unknown proteins
SEQ ID n° 5239	3103	949	SA-1869.1	836176-837213 p	Similar to unknown proteins
SEQ ID n° 5240	3104	950	SA-1870.1	837293-837901 p	manganese-dependent superoxide dismutase
SEQ ID n° 5241	3105	951	SA-1871.1	838239-839090 p	Similar to transcription antiterminator
SEQ ID n° 5242	3106	952	SA-1872.2	839083-840951 p	Similar to beta-glucoside-specific PTS enzyme IIABC
SEQ ID n° 5243	3107	953	SA-1874.1	170247-171179 p	similar to oligopeptide ABC transporter (ATP-binding protein)
SEQ ID n° 5244	3108	954	SA-1875.1	169201-170247 p	similar to oligopeptide ABC transporter (ATP-binding protein)
SEQ ID n° 5245	3109	955	SA-1876.1	168157-169188 p	similar to oligopeptide ABC transporter (permease)
SEQ ID n° 5246	3110	956	SA-1878.1	167233-168147 p	similar to oligopeptide ABC transporter (permease)
SEQ ID n° 5247	3111	957	SA-1879.2	165459-167114 p	similar to oligopeptide ABC transporter (binding protein)

SEQ ID n° 5248	3112	958	SA-188.1	1434490-1435866 m	similar to amino acid transporter
SEQ ID n° 5249	3113	959	SA-1881.2	82132-82437 p	ribosomal protein L24
SEQ ID n° 5250	3114	960	SA-1882.2	82461-83003 p	ribosomal protein L5
SEQ ID n° 5251	3115	961	SA-1883.1	83361-83759 p	ribosomal protein S8
SEQ ID n° 5252	3116	962	SA-1884.1	83869-84405 p	ribosomal protein L6
SEQ ID n° 5253	3117	963	SA-1885.1	84506-84862 p	ribosomal protein L18
SEQ ID n° 5254	3118	964	SA-1886.1	84881-85375 p	ribosomal protein S5
SEQ ID n° 5255	3119	966	SA-1888.1	85390-85569 p	ribosomal protein L30
SEQ ID n° 5256	3120	967	SA-1891.1	85694-86134 p	ribosomal protein L15
SEQ ID n° 5257	3121	968	SA-1892.1	86155-87459 p	similar to preprotein translocase SecY
SEQ ID n° 5258	3122	969	SA-1893.2	87554-88192 p	adenylate kinase
SEQ ID n° 5259	3123	970	SA-1894.2	61615-62718 p	Unknown
SEQ ID n° 5260	3124	971	SA-1897.1	62804-64102 p	similar to adenylosuccinate lyase
SEQ ID n° 5261	3125	972	SA-1898.1	64256-65158 p	similar to unknown proteins
SEQ ID n° 5262	3126	973	SA-1899.1	65447-66445 p	similar to Holliday junction DNA helicase, subunit B
SEQ ID n° 5263	3127	974	SA-190.1	1433541-1434485 m	similar to homocysteine S-methyltransferase
SEQ ID n° 5264	3128	975	SA-1900.3	66597-67034 p	similar to protein-tyrosine phosphatase
SEQ ID n° 5265	3129	976	SA-1901.2	974448-974816 m	similar to transcriptional regulator (GntR family)
SEQ ID n° 5266	3130	977	SA-1902.2	974961-978065 p	DNA polymerase III (alpha subunit)
SEQ ID n° 5267	3131	978	SA-1904.1	978146-979168 p	similar to 6-phosphofructokinase
SEQ ID n° 5268	3132	979	SA-1906.1	979217-980719 p	similar to pyruvate kinase
SEQ ID n° 5269	3133	980	SA-1908.2	980890-981447 p	similar to type-1 signal peptidase
SEQ ID n° 5270	3134	981	SA-1909.2	1996039-1996464 m	Similar to galactose 6-P isomerase (A subunit)
SEQ ID n° 5271	3135	982	SA-191.2	1431339-1433447 p	similar to ATP-dependent Clp proteinase (ATP-binding subunit), ClpL
SEQ ID n° 5272	3136	983	SA-1910.1	1995503-1996018 m	Similar to galactose 6-P isomerase (B subunit)
SEQ ID n° 5273	3137	984	SA-1911.1	1994560-1995492 m	Similar to tagatose-6-phosphate kinase
SEQ ID n° 5274	3138	985	SA-1912.1	1993581-1994558 m	Similar to tagatose 1,6-diP aldolase
SEQ ID n° 5275	3139	986	SA-1913.1	1992631-1993527 m	Similar to unknown proteins
SEQ ID n° 5276	3140	987	SA-1915.1	1991685-1992536 m	Similar to unknown proteins
SEQ ID n° 5277	3141	988	SA-1916.2	1990451-1991584 m	Similar to sugar ABC transporter (ATP-binding protein)
SEQ ID n° 5278	3142	989	SA-1918.3	2104019-2105230 m	similar to phosphopentomutase
SEQ ID n° 5279	3143	990	SA-1919.1	2105297-2105968 m	similar to deoxyribose-phosphate aldolase
SEQ ID n° 5280	3144	991	SA-1921.1	2105998-2107200 m	similar to transport system permease protein
SEQ ID n° 5281	3145	992	SA-1922.1	2107221-2108000 m	similar to uridine phosphorylase
SEQ ID n° 5282	3146	993	SA-1923.1	2108158-2108895 p	similar to transcriptional regulator (GntR family)

SEQ ID n° 5283	3147	994	SA-1924.1	2108908-2109204 p	Unknown
SEQ ID n° 5284	3148	996	SA-1926.2	2109304-2110926 m	chaperonin GroEL
SEQ ID n° 5285	3149	999	SA-1929.2	1348211-1349551 m	similar to multidrug resistance protein
SEQ ID n° 5286	3150	1000	SA-193.2	1430405-1430905 m	Ribosomal protein L10
SEQ ID n° 5287	3151	1001	SA-1930.1	1347385-1348152 p	Similar to unknown proteins
SEQ ID n° 5288	3152	1002	SA-1931.1	1345445-1347226 m	exinuclease ABC (subunit C)
SEQ ID n° 5289	3153	1003	SA-1932.2	1343622-1345403 p	Transmembrane N-terminal domain, C-terminal domain similar to hydrolases
SEQ ID n° 5290	3154	1004	SA-1933.2	1342850-1343452 m	similar to NADH dehydrogenase
SEQ ID n° 5291	3155	1005	SA-1934.2	1341397-1342803 m	similar to dipeptidase
SEQ ID n° 5292	3156	1006	SA-1935.1	1340716-1341300 m	similar to unknown proteins
SEQ ID n° 5293	3157	1007	SA-1937.1	1339967-1340701 m	similar to unknown proteins
SEQ ID n° 5294	3158	1009	SA-194.1	1429976-1430341 m	ribosomal protein L7/L12
SEQ ID n° 5295	3159	1010	SA-1940.3	782771-783706 p	Hpr (ser) kinase/phosphatase
SEQ ID n° 5296	3160	1012	SA-1942.1	782280-782543 p	Similar to unknown proteins
SEQ ID n° 5297	3161	1013	SA-1943.1	781741-782199 p	Similar to unknown proteins
SEQ ID n° 5298	3162	1014	SA-1944.1	779616-781778 p	Similar to unknown proteins
SEQ ID n° 5299	3163	1015	SA-1945.1	778468-779511 p	Similar to oxidoreductase
SEQ ID n° 5300	3164	1016	SA-1946.1	778207-778335 m	Similar to unknown proteins
SEQ ID n° 5301	3165	1017	SA-1949.1	777299-778201 m	Similar to unknown proteins
SEQ ID n° 5302	3166	1020	SA-1951.1	776487-777299 m	Similar to unknown proteins
SEQ ID n° 5303	3167	1021	SA-1952.2	774840-776450 p	Similar to signal recognition particle and to cell division protein FtsY
SEQ ID n° 5304	3168	1022	SA-1953.2	256648-257793 m	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)
SEQ ID n° 5305	3169	1023	SA-1954.1	256013-256648 m	similar to choline ABC transporter (permease)
SEQ ID n° 5306	3170	1024	SA-1955.1	255084-256010 m	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)
SEQ ID n° 5307	3171	1025	SA-1956.1	254433-255083 m	similar to glycine betaine/carnitine/choline ABC transporter (permease)
SEQ ID n° 5308	3172	1026	SA-1957.1	253190-254176 p	similar to efflux protein (truncated)
SEQ ID n° 5309	3173	1027	SA-1958.1	252257-253117 p	similar to transcriptional regulator (Rgg like)
SEQ ID n° 5310	3174	1028	SA-1959.1	251133-251414 p	Unknown
SEQ ID n° 5311	3175	1029	SA-196.1	1429562-1429735 m	Unknown
SEQ ID n° 5312	3176	1030	SA-1960.1	250713-251099 p	Unknown
SEQ ID n° 5313	3177	1031	SA-1961.1	250324-250728 p	pseudogene

SEQ ID n° 5314	3178	1032	SA-1962.2	249517-250104 p	Unknown
SEQ ID n° 5315	3179	1033	SA-1963.2	249110-249517 p	Unknown
SEQ ID n° 5316	3180	1034	SA-1966.1	1393216-1395234 m	similar to ATP-dependent helicase
SEQ ID n° 5317	3181	1035	SA-1967.1	1392522-1392881 m	similar to unknown proteins
SEQ ID n° 5318	3182	1036	SA-1968.1	1392147-1392512 m	similar to unknown proteins
SEQ ID n° 5319	3183	1037	SA-1969.1	1390283-1392160 m	similar to conjugative transposon relaxase
SEQ ID n° 5320	3184	1038	SA-197.1	1428750-1429565 m	similar to plasmid protein (replication protein ?)
SEQ ID n° 5321	3185	1039	SA-1971.2	1389373-1390128 m	similar to lactose repressor
SEQ ID n° 5322	3186	1040	SA-1973.3	866024-866614 m	similar to unknown proteins
SEQ ID n° 5323	3187	1041	SA-1974.2	865181-865939 p	Similar to unknown proteins
SEQ ID n° 5324	3188	1042	SA-1975.2	864963-865178 p	similar to transcriptional regulator
SEQ ID n° 5325	3189	1043	SA-1976.2	864595-864918 p	Unknown
SEQ ID n° 5326	3190	1044	SA-1977.1	863331-864149 p	Similar to UDP-D-glucose:galactosyl glucosyltransferase
SEQ ID n° 5327	3191	1045	SA-1978.1	862527-863243 p	similar to unknown proteins
SEQ ID n° 5328	3192	1046	SA-198.1	1428091-1428513 m	similar to methyl transferase (from transposon)
SEQ ID n° 5329	3193	1047	SA-1982.1	859827-862445 p	Alanyl-tRNA synthetase
SEQ ID n° 5330	3194	1048	SA-1983.2	859326-859811 p	similar to unknown proteins
SEQ ID n° 5331	3195	1049	SA-1985.2	1831275-1831817 m	Similar to unknown proteins
SEQ ID n° 5332	3196	1050	SA-1986.1	1828851-1831190 m	Similar to DNA mismatch repair protein MutS
SEQ ID n° 5333	3197	1051	SA-1987.1	1828347-1828847 m	Similar to unknown proteins
SEQ ID n° 5334	3198	1052	SA-1988.1	1827952-1828266 m	Similar to thioredoxin
SEQ ID n° 5335	3199	1053	SA-1989.1	1827313-1827906 p	Similar to transcriptional regulators
SEQ ID n° 5336	3200	1054	SA-199.1	1427312-1428664 m	3.2 DNA restriction/modification and repair
SEQ ID n° 5337	3201	1055	SA-1990.2	1826012-1827136 p	Similar to A/G-specific adenine glycosylase
SEQ ID n° 5338	3202	1056	SA-1992.2	1866085-1866867 m	Similar to unknown protein
SEQ ID n° 5339	3203	1057	SA-1993.1	1864869-1866041 p	Similar to unknown proteins
SEQ ID n° 5340	3204	1058	SA-1994.1	1864198-1864734 m	Similar to unknown proteins
SEQ ID n° 5341	3205	1059	SA-1995.1	1863567-1864154 m	Similar to other proteins
SEQ ID n° 5342	3206	1060	SA-1996.1	1862715-1863440 m	Similar to transcriptional regulator
SEQ ID n° 5343	3207	1061	SA-1997.1	1861455-1862645 p	Similar to Cyclopropane fatty acid synthase
SEQ ID n° 5344	3208	1062	SA-1998.1	1860995-1861294 p	Similar to unknown proteins
SEQ ID n° 5345	3209	1064	SA-2.1	1025155-1025412 m	Unknown
SEQ ID n° 5346	3210	1065	SA-20.1	1041931-1043013 m	similar to unknown proteins
SEQ ID n° 5347	3211	1066	SA-200.1	1426882-1427328 m	similar to unknown proteins
SEQ ID n° 5348	3212	1067	SA-2000.1	1859666-1860112 m	Similar to unknown proteins
SEQ ID n° 5349	3213	1068	SA-2004.2	1819075-1821903 m	excinuclease ABC (subunit A)

SEQ ID n° 5350	3214	1069	SA-2005.1	1822035-1822706 m	Similar to unknown proteins
SEQ ID n° 5351	3215	1070	SA-2006.1	1822731-1823675 m	Similar to transport protein
SEQ ID n° 5352	3216	1071	SA-2007.1	1823845-1824084 m	30S ribosomal protein S18
SEQ ID n° 5353	3217	1072	SA-2008.2	1824129-1824620 m	Similar to single strand binding protein
SEQ ID n° 5354	3218	1073	SA-2009.2	1824632-1824919 m	30S ribosomal protein S6
SEQ ID n° 5355	3219	1078	SA-2015.2	14009-15985 p	cell division protein FtsH
SEQ ID n° 5356	3220	1079	SA-2017.2	13444-13986 p	similar to hypoxanthine guanine phosphoribosyltransferase
SEQ ID n° 5357	3221	1080	SA-2018.2	12165-13439 p	similar to other protein
SEQ ID n° 5358	3222	1081	SA-2019.3	10877-12163 p	weakly similar to beta-lactamase
SEQ ID n° 5359	3223	1082	SA-202.1	1426503-1426889 m	similar to unknown proteins
SEQ ID n° 5360	3224	1083	SA-2022.2	1569925-1570953 m	Similar to transcriptional regulator (LacI family)
SEQ ID n° 5361	3225	1084	SA-2024.1	1571209-1572456 p	similar to maltose ABC transporter (binding protein)
SEQ ID n° 5362	3226	1085	SA-2025.1	1572554-1573924 p	Similar to maltose/maltodextrin ABC transport system (permease)
SEQ ID n° 5363	3227	1086	SA-2026.1	1573924-1574760 p	similar to maltodextrin ABC transporter (permease)
SEQ ID n° 5364	3228	1087	SA-2028.1	1574991-1576463 p	similar to di-tripeptide ABC transporter (permease)
SEQ ID n° 5365	3229	1088	SA-2029.1	1576508-1576972 m	Similar to putative mutator MutT protein
SEQ ID n° 5366	3230	1089	SA-203.1	1426273-1426506 m	Unknown
SEQ ID n° 5367	3231	1090	SA-2030.2	1577056-1577244 m	Unknown
SEQ ID n° 5368	3232	1092	SA-2033.2	2150270-2150875 p	similar to plasmid replication protein
SEQ ID n° 5369	3233	1093	SA-2034.1	2149459-2150064 p	similar to unknown proteins
SEQ ID n° 5370	3234	1094	SA-2035.4	2148362-2149462 p	similar to other proteins
SEQ ID n° 5371	3235	1095	SA-2037.4	2148042-2148362 p	Unknown
SEQ ID n° 5372	3236	1096	SA-2038.1	2147391-2147840 p	Unknown
SEQ ID n° 5373	3237	1097	SA-2040.1	2145655-2147070 p	similar to unknown proteins
SEQ ID n° 5374	3238	1098	SA-2041.2	2145266-2145604 p	similar to cadmium-efflux system accessory protein
SEQ ID n° 5375	3239	1099	SA-2042.2	2144640-2145254 p	putative transport protein
SEQ ID n° 5376	3240	1100	SA-2045.2	75018-76298 p	hypothetical transport protein
SEQ ID n° 5377	3241	1101	SA-2047.1	73423-74913 p	similar to threonine synthase
SEQ ID n° 5378	3242	1102	SA-2048.1	72286-73302 p	similar to alcohol dehydrogenase
SEQ ID n° 5379	3243	1103	SA-205.1	1425689-1426270 m	Unknown
SEQ ID n° 5380	3244	1104	SA-2050.2	69465-72107 p	similar to alcohol-acetaldehyde dehydrogenase
SEQ ID n° 5381	3245	1105	SA-2051.2	849073-849753 p	similar to unknown proteins
SEQ ID n° 5382	3246	1106	SA-2052.1	847892-848920 m	Similar to S-adenosylmethionine tRNA ribosyltransferase
SEQ ID n° 5383	3247	1107	SA-2053.1	847368-847805 m	Similar to transcriptional regulator, MarR family
SEQ ID n° 5384	3248	1108	SA-2054.1	846249-847313 p	Similar to unknown proteins

SEQ ID n° 5385	3249	1109	SA-2056.1	844892-846148 p	Similar to permease (gluconate ?)
SEQ ID n° 5386	3250	1110	SA-2059.2	843725-844867 p	Similar to putative glycerate kinase
SEQ ID n° 5387	3251	1111	SA-206.1	1425120-1425608 m	similar to unknown proteins
SEQ ID n° 5388	3252	1112	SA-2060.1	842464-843558 m	Similar to unknown proteins
SEQ ID n° 5389	3253	1113	SA-2061.2	840968-842395 p	Similar to beta-glucosidase
SEQ ID n° 5390	3254	1114	SA-2062.2	81399-81659 p	ribosomal protein S17
SEQ ID n° 5391	3255	1115	SA-2063.2	81167-81373 p	ribosomal protein L29
SEQ ID n° 5392	3256	1116	SA-2065.1	80744-81157 p	ribosomal protein L16
SEQ ID n° 5393	3257	1117	SA-2066.1	80087-80740 p	ribosomal protein S3
SEQ ID n° 5394	3258	1118	SA-2067.1	79730-80074 p	ribosomal protein L22
SEQ ID n° 5395	3259	1119	SA-2069.1	79436-79714 p	ribosomal protein S19
SEQ ID n° 5396	3260	1120	SA-207.1	1423303-1425120 m	similar to plasmid transfer complex protein Trsk
SEQ ID n° 5397	3261	1121	SA-2071.1	78504-79337 p	ribosomal protein L2
SEQ ID n° 5398	3262	1122	SA-2072.1	78190-78486 p	ribosomal protein L23
SEQ ID n° 5399	3263	1123	SA-2073.1	77567-78190 p	ribosomal protein L4
SEQ ID n° 5400	3264	1124	SA-2074.2	76917-77543 p	ribosomal protein L3
SEQ ID n° 5401	3265	1125	SA-2075.2	76504-76812 p	ribosomal protein S10
SEQ ID n° 5402	3266	1127	SA-2077.3	1351826-1352680 p	similar to transposase (truncated)
SEQ ID n° 5403	3267	1128	SA-2078.2	1363404-1363694 p	similar to transposase, N-terminal part
SEQ ID n° 5404	3268	1129	SA-2079.2	1362139-1363350 m	Unknown
SEQ ID n° 5405	3269	1130	SA-208.1	1423041-1423283 m	Unknown
SEQ ID n° 5406	3270	1131	SA-2082.3	1358151-1361603 m	streptococcal C5a peptidase
SEQ ID n° 5407	3271	1132	SA-2083.4	1800964-1802403 p	Similar to sucrose-6-phosphate hydrolase
SEQ ID n° 5408	3272	1133	SA-2084.1	1802405-1803367 p	Similar to transcriptional regulator (LacI family)
SEQ ID n° 5409	3273	1134	SA-2085.1	1803454-1803888 m	Similar to transcription termination protein NusB
SEQ ID n° 5410	3274	1135	SA-2086.1	1803881-1804270 m	Similar to unknown proteins
SEQ ID n° 5411	3275	1136	SA-2087.1	1804359-1804919 m	Similar to translation elongation factor EF-P
SEQ ID n° 5412	3276	1137	SA-2090.1	1806098-1807819 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5413	3277	1139	SA-2092.2	1807809-1809566 m	similar to ABC transporter (binding protein)
SEQ ID n° 5414	3278	1140	SA-2095.2	464590-465939 p	glucose-6-phosphate isomerase
SEQ ID n° 5415	3279	1141	SA-2096.1	466261-466788 p	similar to unknown protein
SEQ ID n° 5416	3280	1142	SA-2097.1	466785-467456 p	similar to unknown protein
SEQ ID n° 5417	3281	1143	SA-2099.1	467588-468631 p	similar to putative ABC transporter (binding protein)
SEQ ID n° 5418	3282	1144	SA-21.1	1043024-1043371 m	Unknown
SEQ ID n° 5419	3283	1145	SA-210.1	1422170-1423024 m	similar to unknown protein from conjugative transposon
SEQ ID n° 5420	3284	1146	SA-2100.1	468722-469621 m	similar to UDP-glucose pyrophosphorylase

SEQ ID n° 5421	3285	1147	SA-2101.1	469658-470674 m	similar to glycerol-3-phosphate dehydrogenase
SEQ ID n° 5422	3286	1148	SA-2102.1	470844-471173 p	similar to ribonuclease P protein component
SEQ ID n° 5423	3287	1149	SA-2103.2	471186-472001 p	similar to unknown protein and to B. subtilis SpoIIJ protein
SEQ ID n° 5424	3288	1150	SA-2104.2	2047351-2048703 m	Unknown
SEQ ID n° 5425	3289	1151	SA-2105.1	2048713-2049804 m	Unknown
SEQ ID n° 5426	3290	1152	SA-2106.1	2050044-2050988 m	similar to hypothetical transcriptional regulator
SEQ ID n° 5427	3291	1153	SA-2107.1	2051052-2051441 m	Unknown
SEQ ID n° 5428	3292	1154	SA-2109.1	2052439-2052756 p	Unknown
SEQ ID n° 5429	3293	1155	SA-2111.1	1421755-1422108 m	similar to unknown protein from conjugative transposons
SEQ ID n° 5430	3294	1156	SA-2110.1	2052805-2052933 p	Unknown
SEQ ID n° 5431	3295	1157	SA-2112.1	2053374-2053664 p	Similar to unknown proteins
SEQ ID n° 5432	3296	1158	SA-2113.1	2053763-2054137 p	Unknown
SEQ ID n° 5433	3297	1159	SA-2115.1	2054427-2055155 p	Unknown
SEQ ID n° 5434	3298	1160	SA-2116.1	2055450-2055644 p	Unknown
SEQ ID n° 5435	3299	1161	SA-2117.1	2055763-2056128 m	Unknown
SEQ ID n° 5436	3300	1162	SA-2118.2	2056356-2056754 m	Unknown
SEQ ID n° 5437	3301	1163	SA-2119.2	1523036-1523971 m	Similar to manganese-dependent inorganic pyrophosphatase
SEQ ID n° 5438	3302	1164	SA-2121.1	1419453-1421798 m	similar to plamsid transfer complex protein TraE
SEQ ID n° 5439	3303	1165	SA-2120.1	1524088-1524876 m	Similar to pyruvate-formate lyase activating enzyme
SEQ ID n° 5440	3304	1166	SA-2122.1	1524944-1526278 m	Similar to unknown protein
SEQ ID n° 5441	3305	1167	SA-2123.1	1526458-1527024 m	Similar to unknown proteins
SEQ ID n° 5442	3306	1168	SA-2124.1	1527017-1527952 m	Similar to unknown protein
SEQ ID n° 5443	3307	1169	SA-2125.3	1528045-1528686 m	Similar to unknown protein
SEQ ID n° 5444	3308	1170	SA-2126.3	1528667-1529251 m	similar to unknown protein
SEQ ID n° 5445	3309	1171	SA-2128.2	1352904-1353194 p	similar to transposase
SEQ ID n° 5446	3310	1172	SA-2129.1	1353352-1353732 p	degenerate transposase
SEQ ID n° 5447	3311	1173	SA-2130.1	1353776-1354036 p	degenerate transposase
SEQ ID n° 5448	3312	1174	SA-2131.1	1354082-1354315 p	Degenerate transposase
SEQ ID n° 5449	3313	1175	SA-2132.1	1354514-1356982 m	Similar to histidine triad protein
SEQ ID n° 5450	3314	1176	SA-2133.2	1356995-1357915 m	laminin-binding surface protein
SEQ ID n° 5451	3315	1177	SA-2134.2	625678-626331 p	similar to two-component response regulator VncR
SEQ ID n° 5452	3316	1178	SA-2136.1	626328-627647 p	similar to two-component sensor histidine kinase VncS
SEQ ID n° 5453	3317	1179	SA-2137.1	627699-628358 m	Similar to transposase C-terminal end (truncated)
SEQ ID n° 5454	3318	1180	SA-2138.1	628524-628724 p	Similar to unknown proteins
SEQ ID n° 5455	3319	1182	SA-214.1	1416656-1419451 m	similar to plasmid and conjugative transposon protein
SEQ ID n° 5456	3320	1183	SA-2140.1	629379-630584 p	similar to FtsW and RodA proteins

SEQ ID n° 5457	3321	1184	SA-2141.1	630703-631263 p	similar to unknown protein
SEQ ID n° 5458	3322	1185	SA-2143.2	631264-633216 p	DNA gyrase, subunit B
SEQ ID n° 5459	3323	1186	SA-2145.1	1900897-1901571 m	similar to two-component response regulator
SEQ ID n° 5460	3324	1187	SA-2146.1	1901927-1902061 m	ribosomal protein L34
SEQ ID n° 5461	3325	1188	SA-2147.1	1902245-1903600 m	Similar to unknown proteins
SEQ ID n° 5462	3326	1189	SA-2148.1	1903847-1905574 m	similar to betaine ABC transporter permease and substrate binding protein
SEQ ID n° 5463	3327	1190	SA-2149.2	1905593-1906816 m	similar to glycine betaine ABC transporter (ATP-binding protein)
SEQ ID n° 5464	3328	1191	SA-2156.2	2095537-2096076 m	similar to transcription antitermination factor nusG
SEQ ID n° 5465	3329	1192	SA-2157.1	2094285-2095490 p	similar to glycosyl transferase
SEQ ID n° 5466	3330	1193	SA-2158.1	2093024-2094220 p	similar to glycosyl transferase
SEQ ID n° 5467	3331	1194	SA-2159.1	2091974-2092819 m	similar to unknown protein
SEQ ID n° 5468	3332	1195	SA-216.1	1415704-1416552 p	similar to unknown proteins
SEQ ID n° 5469	3333	1196	SA-2160.2	2090684-2091931 p	similar to transporter
SEQ ID n° 5470	3334	1197	SA-2161.2	1244116-1245459 p	similar to staphylokinase and streptokinase
SEQ ID n° 5471	3335	1198	SA-2162.1	1243011-1243697 m	similar to unknown protein
SEQ ID n° 5472	3336	1199	SA-2163.1	1241896-1242903 m	similar to unknown protein
SEQ ID n° 5473	3337	1200	SA-2165.1	1240446-1241822 m	similar to succinic semialdehyde dehydrogenase
SEQ ID n° 5474	3338	1201	SA-2166.1	1239758-1240303 m	similar to glycine betaine transporter (N-terminal end)
SEQ ID n° 5475	3339	1202	SA-2167.1	1238756-1239757 m	similar to glycine betaine transporter (C-terminal end)
SEQ ID n° 5476	3340	1203	SA-2168.1	1238248-1238733 p	similar to unknown protein C-terminal part
SEQ ID n° 5477	3341	1204	SA-2169.3	1237847-1238407 p	similar to unknown protein (N-terminal part)
SEQ ID n° 5478	3342	1205	SA-217.1	1415117-1415707 p	similar to unknown proteins
SEQ ID n° 5479	3343	1207	SA-2172.2	1100661-1102145 m	similar to carbon starvation protein A
SEQ ID n° 5480	3344	1208	SA-2173.1	1102301-1103035 m	similar to two-component response regulator lytR
SEQ ID n° 5481	3345	1209	SA-2174.1	1103047-1104786 m	similar to two-component sensor histidine kinase LytS
SEQ ID n° 5482	3346	1212	SA-2178.1	1106331-1106708 m	Unknown
SEQ ID n° 5483	3347	1213	SA-2180.3	1107097-1107420 m	Unknown
SEQ ID n° 5484	3348	1214	SA-2182.3	1107768-1108301 m	Unknown
SEQ ID n° 5485	3349	1215	SA-2184.1	825634-827220 p	Similar to ATP-dependent RNA helicase, DEAD-box family
SEQ ID n° 5486	3350	1216	SA-2185.1	824569-825399 p	putative ABC transporter (binding protein)
SEQ ID n° 5487	3351	1217	SA-2186.1	823891-824553 p	similar to ABC transporter (permease)
SEQ ID n° 5488	3352	1218	SA-2187.1	823164-823898 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5489	3353	1219	SA-2188.1	822663-823043 p	Similar to unknown proteins
SEQ ID n° 5490	3354	1220	SA-2190.1	821033-822577 p	Similar to peptide-chain-release factor 3

SEQ ID n° 5491	3355	1221	SA-2192.2	819318-820856 p	Similar to surface proteins, putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5492	3356	1222	SA-2195.3	1140455-1142050 m	similar to cardiolipin synthetase
SEQ ID n° 5493	3357	1223	SA-2196.2	1142169-1143839 m	similar to formate-tetrahydrofolate ligase
SEQ ID n° 5494	3358	1224	SA-2197.2	1143928-1144947 m	similar to lipoate-protein ligase
SEQ ID n° 5495	3359	1225	SA-2198.2	1144974-1145852 m	similar to unknown protein
SEQ ID n° 5496	3360	1226	SA-22.1	1043463-1045295 m	similar to ATP-dependent Clp protease ATP-binding subunit ClpA
SEQ ID n° 5497	3361	1227	SA-220.1	1410000-1414904 m	Putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5498	3362	1228	SA-2200.2	120068-121351 p	similar to trigger factor (prolyl isomerase)
SEQ ID n° 5499	3363	1229	SA-2201.1	119069-119911 m	similar to unknown protein
SEQ ID n° 5500	3364	1230	SA-2202.1	118463-119032 p	similar to unknown protein
SEQ ID n° 5501	3365	1231	SA-2203.1	118002-118466 p	similar to unknown protein
SEQ ID n° 5502	3366	1232	SA-2204.1	117234-117992 p	similar to phosphomethylpyrimidine kinase
SEQ ID n° 5503	3367	1233	SA-2205.1	116495-117271 p	similar to tRNA pseudouridine synthase A
SEQ ID n° 5504	3368	1234	SA-2206.2	115177-116424 p	Similar to other proteins
SEQ ID n° 5505	3369	1235	SA-2207.2	573586-574413 p	similar to unknown proteins
SEQ ID n° 5506	3370	1236	SA-2208.2	574400-574873 p	similar to transcriptional regulator
SEQ ID n° 5507	3371	1237	SA-2210.1	574885-576543 p	similar to DNA repair and recombination protein RecN
SEQ ID n° 5508	3372	1238	SA-2212.1	576656-577492 p	similar to unknown proteins
SEQ ID n° 5509	3373	1239	SA-2213.1	577467-578324 p	similar to unknown proteins
SEQ ID n° 5510	3374	1240	SA-2214.2	578299-578901 p	Similar to unknown proteins
SEQ ID n° 5511	3375	1242	SA-2216.2	1250993-1251574 m	similar to unknown proteins
SEQ ID n° 5512	3376	1243	SA-2217.1	1250789-1250983 m	similar to unknown protein
SEQ ID n° 5513	3377	1244	SA-2219.1	1250191-1250733 m	similar to unknown protein
SEQ ID n° 5514	3378	1245	SA-222.1	1409808-1409999 m	Unknown
SEQ ID n° 5515	3379	1246	SA-2220.1	1249935-1250132 m	similar to unknown protein
SEQ ID n° 5516	3380	1247	SA-2221.1	1249362-1249913 m	similar to unknown protein
SEQ ID n° 5517	3381	1248	SA-2222.1	1248557-1249198 m	similar to transcriptional regulator, GntR family
SEQ ID n° 5518	3382	1249	SA-2223.1	1247788-1248552 m	similar to unknown protein
SEQ ID n° 5519	3383	1250	SA-2224.1	1247129-1247788 m	similar to other protein
SEQ ID n° 5520	3384	1251	SA-2225.1	1246562-1247056 m	similar to thiol peroxidase
SEQ ID n° 5521	3385	1252	SA-2226.2	1245953-1246486 p	similar to unknown proteins
SEQ ID n° 5522	3386	1253	SA-2227.2	2001195-2001659 m	Similar to PTS enzyme IIA
SEQ ID n° 5523	3387	1254	SA-2228.1	2000888-2001193 m	Similar to PTS enzyme IIB
SEQ ID n° 5524	3388	1255	SA-2231.1	1999400-2000848 m	Similar to galactitol-specific PTS enzyme IIC

SEQ ID n° 5525	3389	1256	SA-2232.2	1996712-1999162 m	Similar to neuraminidase
SEQ ID n° 5526	3390	1257	SA-2233.2	1544787-1545722 m	Similar to putative rhamnosyltransferase
SEQ ID n° 5527	3391	1258	SA-2235.1	1545724-1546779 m	Similar to nucleotide-sugar dehydratase
SEQ ID n° 5528	3392	1259	SA-2236.1	1546782-1547501 m	Similar to 4-diphosphocytidyl-2C-methyl-D-erythritol synthase
SEQ ID n° 5529	3393	1260	SA-2237.1	1547504-1548328 m	Similar to Pneumococcal LicD2 protein involved in phosphorylcholine metabolism
SEQ ID n° 5530	3394	1261	SA-2238.1	1548353-1550086 m	possible surface protein
SEQ ID n° 5531	3395	1262	SA-2239.1	1550079-1550432 m	Similar to unknown protein
SEQ ID n° 5532	3396	1263	SA-224.1	1409273-1409824 m	Unknown
SEQ ID n° 5533	3397	1264	SA-2240.1	1550429-1551157 m	Similar to putative glycosyl transferase
SEQ ID n° 5534	3398	1265	SA-2241.4	1551162-1552103 m	similar to rhamnosyltransferase
SEQ ID n° 5535	3399	1266	SA-2242.3	145779-146144 p	similar to unknown protein
SEQ ID n° 5536	3400	1267	SA-2244.3	146144-147808 p	similar to unknown protein
SEQ ID n° 5537	3401	1268	SA-2245.1	147956-148840 p	similar to unknown protein
SEQ ID n° 5538	3402	1270	SA-2247.1	149922-150662 m	similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 5539	3403	1271	SA-2248.3	150672-152222 m	similar to amino acid (glutamine) ABC transporter (binding protein)
SEQ ID n° 5540	3404	1272	SA-225.1	1408629-1409222 m	Unknown
SEQ ID n° 5541	3405	1275	SA-2253.4	1076528-1077445 m	similar to phosphate ABC transporter (permease)
SEQ ID n° 5542	3406	1276	SA-2254.2	1075651-1076538 m	similar to phosphate ABC transporter (permease)
SEQ ID n° 5543	3407	1277	SA-2255.1	1074836-1075639 m	similar to phosphate ABC transporter (ATP-binding protein)
SEQ ID n° 5544	3408	1278	SA-2256.1	1074066-1074824 m	similar to phosphate ABC transporter (ATP-binding protein)
SEQ ID n° 5545	3409	1279	SA-2258.2	1073379-1074032 m	similar to phosphate uptake regulatory protein
SEQ ID n° 5546	3410	1280	SA-2259.2	1086749-1089379 m	similar to hypothetical ABC transporter (permease)
SEQ ID n° 5547	3411	1281	SA-2260.1	1089391-1090092 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5548	3412	1282	SA-2261.3	1090229-1092349 m	similar to DNA topoisomerase I
SEQ ID n° 5549	3413	1284	SA-2264.3	176720-177571 p	similar to 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
SEQ ID n° 5550	3414	1285	SA-2265.2	177657-178100 p	similar to transcriptional regulator
SEQ ID n° 5551	3415	1286	SA-2266.2	178103-178813 p	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5552	3416	1287	SA-2267.1	178803-179615 p	Similar to (metal) ABC transporter (permease)
SEQ ID n° 5553	3417	1288	SA-2268.2	179827-180858 m	similar to DNase (competence related)
SEQ ID n° 5554	3418	1289	SA-2269.2	1617901-1619289 p	similar to amino acid permease
SEQ ID n° 5555	3419	1290	SA-2271.1	1619373-1619840 m	Similar to small protein SmpB
SEQ ID n° 5556	3420	1291	SA-2272.2	1619843-1622248 m	similar to exoribonuclease R
SEQ ID n° 5557	3421	1292	SA-2274.2	2111481-2112284 m	Similar to ABC transporter (ATP-binding protein)

SEQ ID n° 5558	3422	1293	SA-2276.1	2112289-2113185 m	Similar to ABC transporter (permease)
SEQ ID n° 5559	3423	1294	SA-2278.1	2113201-2114163 m	putative ABC transporter (binding protein)
SEQ ID n° 5560	3424	1295	SA-2279.1	2114628-2115425 m	similar to unknown proteins
SEQ ID n° 5561	3425	1296	SA-2280.1	2115611-2116471 p	similar to unknown proteins
SEQ ID n° 5562	3426	1297	SA-2281.2	2116514-2117245 p	similar to unknown proteins
SEQ ID n° 5563	3427	1298	SA-2282.2	1065700-1066194 m	Unknown
SEQ ID n° 5564	3428	1299	SA-2283.1	1066231-1066419 m	Unknown
SEQ ID n° 5565	3429	1300	SA-2285.1	1066541-1068106 m	similar to signal recognition particle chain Ffh
SEQ ID n° 5566	3430	1301	SA-2287.1	1068124-1068456 m	similar to unknown protein
SEQ ID n° 5567	3431	1302	SA-2288.2	1068545-1069858 m	similar to two-component sensor histidine kinase
SEQ ID n° 5568	3432	1303	SA-2289.2	827664-828428 m	Similar to unknown proteins
SEQ ID n° 5569	3433	1304	SA-2291.1	828563-829303 p	Similar to other proteins
SEQ ID n° 5570	3434	1305	SA-2292.1	829403-830056 p	Similar to competence protein ComEA
SEQ ID n° 5571	3435	1307	SA-2294.2	830049-832277 p	similar to competence protein CelB (ComEC)
SEQ ID n° 5572	3436	1308	SA-2295.2	832403-833212 p	Similar to unknown proteins
SEQ ID n° 5573	3437	1309	SA-2296.2	1540196-1541044 m	Similar to rhamnosyl transferase I
SEQ ID n° 5574	3438	1310	SA-2297.1	1539067-1540206 m	Similar to putative hexosyltransferase
SEQ ID n° 5575	3439	1311	SA-2298.1	1537486-1539015 p	Similar to transcription regulator
SEQ ID n° 5576	3440	1312	SA-23.1	1045297-1047045 m	similar to plasmid DNA topoisomerase
SEQ ID n° 5577	3441	1313	SA-230.1	1402425-1408625 m	similar to putative helicase and methylase
SEQ ID n° 5578	3442	1314	SA-2300.3	1534497-1537202 m	Similar to fibronectin binding protein, peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5579	3443	1315	SA-2302.2	1095272-1096246 m	similar to ferrichrome ABC transporter (permease)
SEQ ID n° 5580	3444	1316	SA-2303.1	1096243-1097205 m	similar to (iron?) ABC transporter (permease)
SEQ ID n° 5581	3445	1317	SA-2304.1	1097444-1097992 m	similar to sugar O-acetyltransferase
SEQ ID n° 5582	3446	1318	SA-2305.1	1098013-1098774 m	similar to ribonuclease HII
SEQ ID n° 5583	3447	1319	SA-2306.1	1098761-1099612 m	similar to unknown protein, hypothetical GTP binding protein
SEQ ID n° 5584	3448	1320	SA-2307.2	1099888-1100460 m	similar to unknown proteins
SEQ ID n° 5585	3449	1321	SA-2308.2	852521-853672 p	similar to multidrug resistance protein
SEQ ID n° 5586	3450	1322	SA-2309.1	853724-854671 p	similar to hypothetical transcription factor
SEQ ID n° 5587	3451	1323	SA-231.1	1402054-1402353 m	Unknown
SEQ ID n° 5588	3452	1324	SA-2311.1	854687-856492 p	group B oligopeptidase PepB
SEQ ID n° 5589	3453	1325	SA-2312.1	856687-857313 p	similar to hypothetical phosphoglycolate phosphatase
SEQ ID n° 5590	3454	1326	SA-2313.1	857387-858094 p	similar to methyltransferase
SEQ ID n° 5591	3455	1327	SA-2314.2	858155-859084 p	similar to protease maturation protein

SEQ ID n° 5592	3456	1328	SA-2315.2	2119955-2122153 m	similar to anaerobic (class III) ribonucleotide reductase large subunit chain
SEQ ID n° 5593	3457	1329	SA-2316.1	2119737-2119880 m	Unknown
SEQ ID n° 5594	3458	1330	SA-2317.1	2118792-2119724 m	similar to oxidoreductase
SEQ ID n° 5595	3459	1331	SA-2318.1	2118292-2118783 m	similar to unknown proteins
SEQ ID n° 5596	3460	1332	SA-2319.2	2117602-2118219 m	similar to anaerobic ribonucleotide reductase activator
SEQ ID n° 5597	3461	1333	SA-232.1	1401744-1402043 m	Unknown
SEQ ID n° 5598	3462	1334	SA-2322.2	1501840-1503189 m	Similar to glutathione reductase
SEQ ID n° 5599	3463	1335	SA-2324.1	1501345-1501797 p	Similar to unknown protein (secreted protein)
SEQ ID n° 5600	3464	1336	SA-2326.1	1499953-1501098 m	Similar to iron-sulfur cofactor synthesis protein (NifS like)
SEQ ID n° 5601	3465	1337	SA-2327.2	1498737-1499951 m	Similar to hypothetical thiamine biosynthesis protein Thil
SEQ ID n° 5602	3466	1338	SA-2329.3	1497457-1498635 m	similar to capsular polyglutamate biosynthesis
SEQ ID n° 5603	3467	1339	SA-233.1	1400941-1401579 m	Unknown
SEQ ID n° 5604	3468	1340	SA-2330.2	143337-144218 p	similar to fructose-bisphosphate aldolase class-II
SEQ ID n° 5605	3469	1341	SA-2331.2	144303-145220 m	similar to L-2-hydroxyisocaproate dehydrogenase
SEQ ID n° 5606	3470	1342	SA-2332.2	145459-145647 p	50S ribosomal protein L28
SEQ ID n° 5607	3471	1343	SA-2334.1	1532350-1534374 m	putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5608	3472	1344	SA-2335.2	1531360-1532238 m	Similar to sortase protein
SEQ ID n° 5609	3473	1345	SA-2336.1	1530476-1531360 m	Similar to sortase protein
SEQ ID n° 5610	3474	1346	SA-2337.2	1529553-1530479 m	Putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 5611	3475	1347	SA-2339.2	1289559-1290257 m	capsular polysaccharide chain length regulator/exporter
SEQ ID n° 5612	3476	1348	SA-234.1	1399815-1400900 m	similar to unknown proteins
SEQ ID n° 5613	3477	1349	SA-2340.2	1288158-1289546 m	similar to glucose-1-phosphate transferase
SEQ ID n° 5614	3478	1350	SA-2341.1	1287685-1288134 m	beta-1,4-galactosyltransferase enhancer
SEQ ID n° 5615	3479	1351	SA-2342.1	1287212-1287685 m	beta-1,4-galactosyltransferase
SEQ ID n° 5616	3480	1352	SA-2343.1	1286070-1287215 m	capsular polysaccharide repeating-unit polymerase
SEQ ID n° 5617	3481	1353	SA-2344.1	1285105-1286073 m	similar to glycosyl transferase
SEQ ID n° 5618	3482	1354	SA-2345.1	1284124-1285071 m	similar to glycosyl transferase
SEQ ID n° 5619	3483	1355	SA-2346.2	1283084-1284040 m	capsular polysaccharide biosynthesis protein
SEQ ID n° 5620	3484	1357	SA-2348.1	1217890-1218570 m	similar to DNA repair protein RadC
SEQ ID n° 5621	3485	1358	SA-2349.1	1217239-1217877 p	similar to unknown protein
SEQ ID n° 5622	3486	1359	SA-235.1	1399530-1399760 m	Unknown
SEQ ID n° 5623	3487	1360	SA-2350.1	1216737-1217084 p	similar to unknown protein
SEQ ID n° 5624	3488	1361	SA-2351.2	1215620-1216735 p	similar to iron-sulfur cofactor synthesis protein or cysteine desulfurase
SEQ ID n° 5625	3489	1364	SA-2354.1	1552093-1553247 m	similar to putative rhamnosyltransferase

SEQ ID n° 5626	3490	1365	SA-2355.1	1553364-1554218 m	dTDP-L-rhamnose synthase
SEQ ID n° 5627	3491	1366	SA-2356.1	1554308-1554649 m	similar to unknown protein
SEQ ID n° 5628	3492	1367	SA-2357.1	1554758-1555867 m	RNA polymerase major sigma factor RpoD
SEQ ID n° 5629	3493	1368	SA-2358.2	1555875-1557683 m	DNA primase
SEQ ID n° 5630	3494	1369	SA-2359.3	121540-122115 p	similar to DNA-directed RNA polymerase (delta subunit)
SEQ ID n° 5631	3495	1370	SA-236.1	1399144-1399533 m	Unknown
SEQ ID n° 5632	3496	1371	SA-2360.1	122388-123992 p	similar to CTP synthetase
SEQ ID n° 5633	3497	1372	SA-2361.3	124101-125027 p	similar to unknown protein
SEQ ID n° 5634	3498	1374	SA-2364.2	1085765-1086643 m	similar to unknown proteins
SEQ ID n° 5635	3499	1375	SA-2365.1	1085068-1085658 m	Unknown
SEQ ID n° 5636	3500	1376	SA-2366.1	1083763-1085034 m	similar to unknown protein
SEQ ID n° 5637	3501	1377	SA-2367.1	1083319-1083750 m	similar to unknown protein
SEQ ID n° 5638	3502	1378	SA-2368.1	1082343-1083227 m	similar to tRNA pseudouridine 55 synthase
SEQ ID n° 5639	3503	1379	SA-2369.2	1081398-1082330 m	macrolide-efflux protein
SEQ ID n° 5640	3504	1380	SA-237.1	1398807-1399097 m	Unknown
SEQ ID n° 5641	3505	1381	SA-2370.3	1723907-1725709 m	Similar to other proteins
SEQ ID n° 5642	3506	1382	SA-2371.1	1723352-1723834 m	Similar to transcription elongation factor GreA
SEQ ID n° 5643	3507	1383	SA-2372.2	1721788-1723248 m	Similar to putative amidase
SEQ ID n° 5644	3508	1384	SA-2373.2	1560344-1561387 m	similar to ammonium transporter
SEQ ID n° 5645	3509	1385	SA-2374.1	1561467-1562594 p	Similar to unknown protein
SEQ ID n° 5646	3510	1386	SA-2376.1	1562631-1563617 m	Similar to unknown protein
SEQ ID n° 5647	3511	1387	SA-2377.1	1563780-1564085 p	Similar to unknown protein
SEQ ID n° 5648	3512	1388	SA-2378.2	1564243-1565616 m	Similar to glycerol (sugar)-3-phosphate transporter
SEQ ID n° 5649	3513	1389	SA-238.1	1398261-1398737 m	similar to phage repressor-like protein
SEQ ID n° 5650	3514	1390	SA-2380.2	890576-891346 p	similar to hydroxyethyl thiazole kinase (ThiM)
SEQ ID n° 5651	3515	1391	SA-2381.1	889777-890574 p	similar to phosphomethylpyrimidine kinase (ThiD)
SEQ ID n° 5652	3516	1392	SA-2382.1	889096-889752 p	Similar to putative thiamin biosynthesis protein
SEQ ID n° 5653	3517	1394	SA-2384.2	888320-889012 p	Similar to unknown proteins
SEQ ID n° 5654	3518	1395	SA-2386.1	1541041-1542465 m	Similar to capsular polysaccharide synthesis protein
SEQ ID n° 5655	3519	1396	SA-2387.1	1542465-1543829 m	Similar to unknown proteins
SEQ ID n° 5656	3520	1397	SA-2388.2	1543831-1544778 m	Similar to putative rhamnosyltransferase
SEQ ID n° 5657	3521	1398	SA-239.1	1397491-1398261 m	similar to unknown proteins
SEQ ID n° 5658	3522	1399	SA-2390.2	1458427-1460499 p	similar to 5 -nucleotidase, putative peptidoglycan bound protein (LPXTN motif)
SEQ ID n° 5659	3523	1400	SA-2391.1	1460536-1460946 m	similar to polypeptide deformylase
SEQ ID n° 5660	3524	1401	SA-2392.2	1461016-1462365 m	similar to NADP-specific glutamate dehydrogenase

SEQ ID n° 5661	3525	1402	SA-2394.2	1133801-1135033 m	similar to unknown protein
SEQ ID n° 5662	3526	1403	SA-2395.2	1135073-1136614 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5663	3527	1404	SA-2396.2	1480876-1481946 m	similar to other protein, putative transmembrane protein
SEQ ID n° 5664	3528	1405	SA-2397.1	1482083-1483075 m	similar to thioredoxin reductase
SEQ ID n° 5665	3529	1406	SA-2398.1	1483056-1483808 m	similar to tRNA (guanine-N1)-methyltransferase
SEQ ID n° 5666	3530	1407	SA-2399.3	1483798-1484316 m	similar to 16S rRNA processing protein RimM
SEQ ID n° 5667	3531	1408	SA-240.2	1395238-1397478 m	similar to unknown proteins
SEQ ID n° 5668	3532	1409	SA-2404.2	1798860-1800779 m	Similar to sucrose-specific PTS enzyme IIABC
SEQ ID n° 5669	3533	1410	SA-2405.2	515623-516378 m	Similar to methyltransferase
SEQ ID n° 5670	3534	1411	SA-2406.1	515264-515602 m	similar to unknown proteins
SEQ ID n° 5671	3535	1412	SA-2407.1	514957-515262 m	similar to unknown proteins
SEQ ID n° 5672	3536	1413	SA-2409.2	514161-514739 p	similar to integrase (C-terminal part)
SEQ ID n° 5673	3537	1414	SA-241.2	319054-319572 p	similar to unknown proteins
SEQ ID n° 5674	3538	1415	SA-2410.1	513604-514038 p	similar to integrase (N-terminal part)
SEQ ID n° 5675	3539	1416	SA-2411.1	513228-513554 p	Similar to unknown phage proteins
SEQ ID n° 5676	3540	1417	SA-2412.1	512516-513055 p	Unknown
SEQ ID n° 5677	3541	1418	SA-2414.1	511516-512277 p	Similar to plasmid related proteins, Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5678	3542	1419	SA-2415.2	510934-511491 p	Unknown
SEQ ID n° 5679	3543	1420	SA-2416.2	510521-510934 p	Unknown
SEQ ID n° 5680	3544	1421	SA-2418.2	1069842-1070522 m	similar to two-component response regulator
SEQ ID n° 5681	3545	1422	SA-2419.2	1070684-1073233 m	similar to lysyl-aminopeptidase
SEQ ID n° 5682	3546	1423	SA-242.2	318380-318979 m	similar to recombination protein U (RecU)
SEQ ID n° 5683	3547	1424	SA-2420.2	1496953-1497267 m	50S ribosomal protein L21
SEQ ID n° 5684	3548	1426	SA-2422.2	1496608-1496946 m	Similar to unknown proteins
SEQ ID n° 5685	3549	1427	SA-2423.1	1496293-1496586 m	50S ribosomal protein L27
SEQ ID n° 5686	3550	1428	SA-2424.1	1495172-1496077 m	Similar to transcriptional regulator (LysR/MarR family)
SEQ ID n° 5687	3551	1429	SA-2425.1	1494699-1495163 m	Similar to prolipoprotein signal peptidase
SEQ ID n° 5688	3552	1430	SA-2426.3	1493825-1494715 m	Similar to ribosomal large subunit pseudouridine synthase B (hypothetical)
SEQ ID n° 5689	3553	1431	SA-2428.3	1880206-1882284 m	translation elongation factor G
SEQ ID n° 5690	3554	1432	SA-2429.2	1265526-1266473 p	similar to unknown proteins
SEQ ID n° 5691	3555	1433	SA-2430.1	1266592-1267668 p	similar to bacteriophage integrase
SEQ ID n° 5692	3556	1434	SA-2431.3	1267848-1269050 m	similar to ribosomal protein S1-like DNA-binding protein
SEQ ID n° 5693	3557	1436	SA-2433.2	1942754-1943215 p	Similar to other proteins
SEQ ID n° 5694	3558	1437	SA-2434.4	1943374-1944414 m	translation elongation factor EF-Ts

SEQ ID n° 5695	3559	1438	SA-2435.4	1944508-1945278 m	ribosomal protein S2
SEQ ID n° 5696	3560	1440	SA-2437.2	1290268-1290960 m	putative chain length regulator CpsC
SEQ ID n° 5697	3561	1441	SA-2438.1	1290969-1291700 m	Unknown
SEQ ID n° 5698	3562	1443	SA-244.2	316147-318384 m	similar to penicillin-binding protein 1A
SEQ ID n° 5699	3563	1444	SA-2440.3	1291706-1293163 m	Unknown
SEQ ID n° 5700	3564	1446	SA-2442.1	2153854-2154309 p	similar to other proteins (C-terminal end)
SEQ ID n° 5701	3565	1447	SA-2443.1	2152902-2153897 p	Unknown
SEQ ID n° 5702	3566	1448	SA-2445.2	2151206-2152471 p	similar to integrase
SEQ ID n° 5703	3567	1449	SA-2446.2	1369720-1370286 m	similar to plasmid replication protein
SEQ ID n° 5704	3568	1450	SA-2447.1	1369483-1369686 m	Unknown
SEQ ID n° 5705	3569	1451	SA-2448.1	1368284-1369465 m	Similar to integrase (phage-related protein)
SEQ ID n° 5706	3570	1452	SA-2450.2	1206681-1207373 p	Similar to unknown protein
SEQ ID n° 5707	3571	1454	SA-2453.2	849904-850605 p	similar to N-acetylglucosamine-6-phosphate isomerase
SEQ ID n° 5708	3572	1455	SA-2454.1	850677-851633 m	similar to unknown protein
SEQ ID n° 5709	3573	1456	SA-2455.2	851729-852448 p	similar to 16S pseudouridylate synthase
SEQ ID n° 5710	3574	1457	SA-2457.3	67418-69196 p	similar to unknown proteins
SEQ ID n° 5711	3575	1458	SA-2458.3	67041-67421 p	similar to unknown proteins
SEQ ID n° 5712	3576	1459	SA-2459.3	1715970-1716491 m	Similar to unknown proteins
SEQ ID n° 5713	3577	1460	SA-246.1	314767-316101 p	similar to cysteine aminopeptidase C
SEQ ID n° 5714	3578	1461	SA-2460.1	1716559-1717248 m	Similar to unknown proteins
SEQ ID n° 5715	3579	1462	SA-2461.1	1717271-1717774 m	Similar to unknown proteins
SEQ ID n° 5716	3580	1463	SA-2462.1	1717812-1718555 m	Similar to hypothetical rRNA methylase
SEQ ID n° 5717	3581	1464	SA-2463.1	1718587-1718865 p	similar to acylphosphatase
SEQ ID n° 5718	3582	1465	SA-2464.2	1718950-1719882 p	Similar to unknown proteins
SEQ ID n° 5719	3583	1467	SA-2466.2	715787-716482 m	Similar to amino acid ABC transporter (permease)
SEQ ID n° 5720	3584	1468	SA-2467.1	715125-715775 m	similar to ABC transporter (permease)
SEQ ID n° 5721	3585	1469	SA-2468.2	714510-715076 p	similar to unknown proteins
SEQ ID n° 5722	3586	1470	SA-2469.2	713324-714343 p	similar to unknown proteins
SEQ ID n° 5723	3587	1471	SA-247.1	313833-314654 p	similar to NH3-dependent NAD+ synthetase
SEQ ID n° 5724	3588	1472	SA-2470.2	785354-785629 m	Similar to unknown proteins
SEQ ID n° 5725	3589	1473	SA-2471.1	784882-785313 p	Similar to unknown proteins
SEQ ID n° 5726	3590	1474	SA-2472.1	784487-784885 p	Similar to unknown proteins
SEQ ID n° 5727	3591	1475	SA-2473.3	783699-784472 p	similar to prolipoprotein diacylglycerol transferase
SEQ ID n° 5728	3592	1479	SA-2477.2	1492009-1493079 m	Similar to carbamoyl-phosphate synthase, small subunit
SEQ ID n° 5729	3593	1480	SA-2479.2	1490355-1491953 m	similar to carbamoyl-phosphate synthase
SEQ ID n° 5730	3594	1481	SA-248.1	312376-313836 p	similar to unknown proteins

SEQ ID n° 5731	3595	1482	SA-2480.2	1094514-1095275 m	similar to ferrichrome ABC transporter (ATP-binding protein)
SEQ ID n° 5732	3596	1483	SA-2482.2	1093424-1094452 m	similar to ferrichrome ABC transporter (binding protein)
SEQ ID n° 5733	3597	1484	SA-2483.2	1092444-1093286 m	similar to DNA processing Smf protein
SEQ ID n° 5734	3598	1485	SA-2486.2	1614454-1615296 m	similar to oxidoreductase
SEQ ID n° 5735	3599	1486	SA-2488.3	1488909-1490153 m	similar to unknown protein
SEQ ID n° 5736	3600	1487	SA-249.1	311304-312218 p	similar to thioredoxin reductase
SEQ ID n° 5737	3601	1488	SA-2490.2	1488196-1488906 m	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5738	3602	1489	SA-2491.2	1486934-1488181 m	similar to unknown protein
SEQ ID n° 5739	3603	1490	SA-2492.2	981705-983519 p	similar to glucosamine-fructose-6-phosphate aminotransferase
SEQ ID n° 5740	3604	1491	SA-2493.2	1137253-1138329 m	similar to aspartate-semialdehyde dehydrogenase
SEQ ID n° 5741	3605	1493	SA-2495.1	1138538-1139770 m	highly repetitive peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5742	3606	1494	SA-2497.2	1080942-1081355 m	similar to unknown protein
SEQ ID n° 5743	3607	1495	SA-2498.2	1080668-1080949 m	similar to unknown protein
SEQ ID n° 5744	3608	1496	SA-2499.3	1079914-1080678 m	similar to myo-inositol monophosphatase
SEQ ID n° 5745	3609	1497	SA-25.1	1047228-1051958 m	similar to plasmid proteins
SEQ ID n° 5746	3610	1498	SA-250.1	311011-311235 p	similar to unknown proteins
SEQ ID n° 5747	3611	1499	SA-2500.3	1078536-1079846 m	conserved protein
SEQ ID n° 5748	3612	1500	SA-2501.3	1704700-1706052 m	similar to amino acid permease
SEQ ID n° 5749	3613	1501	SA-2502.2	1703979-1704527 m	Similar to rRNA methylase
SEQ ID n° 5750	3614	1502	SA-2503.2	1702674-1703843 m	similar to aminotransferase
SEQ ID n° 5751	3615	1503	SA-2504.2	1701935-1702564 m	uracil phosphoribosyltransferase
SEQ ID n° 5752	3616	1505	SA-2507.2	35466-35705 p	similar to acyl carrier protein
SEQ ID n° 5753	3617	1506	SA-2508.1	34463-35455 p	similar to fatty acid/phospholipid synthesis protein
SEQ ID n° 5754	3618	1507	SA-2509.2	33507-34385 p	similar to unknown transmembrane protein
SEQ ID n° 5755	3619	1508	SA-251.1	310146-310889 p	Similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 5756	3620	1509	SA-2510.2	1207444-1209444 m	similar to cation (K+) transport protein
SEQ ID n° 5757	3621	1510	SA-2511.1	1209579-1210340 m	Similar to oxidoreductase
SEQ ID n° 5758	3622	1511	SA-2512.2	1210402-1211394 m	phosphotransacetylase
SEQ ID n° 5759	3623	1512	SA-2513.2	329975-330604 p	similar to guanylate kinase
SEQ ID n° 5760	3624	1513	SA-2514.1	329104-329808 p	similar to unknown proteins
SEQ ID n° 5761	3625	1514	SA-2515.2	328464-329054 p	Similar to two-component response regulator
SEQ ID n° 5762	3626	1515	SA-2516.2	1615563-1616237 p	similar to putative NAD(P)H-flavin oxidoreductase
SEQ ID n° 5763	3627	1516	SA-2517.1	1616325-1616717 p	similar to S-D-lactoylglutathione methylglyoxal lyase
SEQ ID n° 5764	3628	1517	SA-2518.1	1616846-1617772 p	similar to putative sugar transferase
SEQ ID n° 5765	3629	1519	SA-252.1	309343-310146 p	Similar to amino acid ABC transporter (permease)

SEQ ID n° 5766	3630	1520	SA-2520.3	152358-154232 p	similar to unknown proteins
SEQ ID n° 5767	3631	1521	SA-2522.2	1486533-1486805 m	30S ribosomal protein S16
SEQ ID n° 5768	3632	1522	SA-2523.2	1486281-1486523 m	similar to unknown protein
SEQ ID n° 5769	3633	1523	SA-2524.2	1484638-1486149 m	similar to transcription regulator RofA related
SEQ ID n° 5770	3634	1524	SA-2525.2	960415-962247 p	similar to GTP binding proteins
SEQ ID n° 5771	3635	1525	SA-2526.2	1364617-1366431 m	Exonuclease motif predicted by PFAM
SEQ ID n° 5772	3636	1527	SA-253.1	308436-309248 p	Similar to amino acids ABC transporter (binding protein)
SEQ ID n° 5773	3637	1528	SA-2531.3	1366847-1368082 p	similar to a DNA polymerase like protein from Plasmodium falciparum
SEQ ID n° 5774	3638	1529	SA-2532.1	1062952-1063155 m	Unknown
SEQ ID n° 5775	3639	1530	SA-2533.1	649127-649315 p	Unknown
SEQ ID n° 5776	3640	1531	SA-2534.1	649051-649233 p	pseudogene
SEQ ID n° 5777	3641	1532	SA-2535.1	579004-579279 p	HU like DNA-binding protein
SEQ ID n° 5778	3642	1533	SA-2536.1	572721-573593 p	similar to geranyltransferase (farnesyl diphosphate synthase)
SEQ ID n° 5779	3643	1534	SA-2537.2	565465-565695 p	similar to unknown proteins
SEQ ID n° 5780	3644		SA-2541.2	2046191-2046337 m	Unknown
SEQ ID n° 5781	3645	1541	SA-2548.1	1263082-1263228 m	similar to unknown proteins
SEQ ID n° 5782	3646	1542	SA-2549.1	1264885-1265175 p	Unknown
SEQ ID n° 5783	3647	1543	SA-255.1	306955-308298 p	similar to ATP-dependent RNA helicase
SEQ ID n° 5784	3648	1546	SA-2555.2	507784-508122 p	Unknown
SEQ ID n° 5785	3649	1547	SA-2556.1	1622361-1622597 m	Similar to protein-export protein SecE
SEQ ID n° 5786	3650	1549	SA-2558.1	1631019-1631249 m	Unknown
SEQ ID n° 5787	3651	1550	SA-2559.1	659421-659609 m	similar to transposase
SEQ ID n° 5788	3652	1551	SA-256.1	305847-306857 p	similar to phospho-N-acetylmuramoyl-pentapeptide transferase
SEQ ID n° 5789	3653	1552	SA-2560.1	660639-660797 p	similar to unknown proteins
SEQ ID n° 5790	3654	1559	SA-257.1	303587-305845 p	similar to penicillin-binding protein 2X
SEQ ID n° 5791	3655	1560	SA-2571.3	984145-984786 p	Similar to ABC transporter (permease)
SEQ ID n° 5792	3656	1561	SA-2572.1	983681-984010 p	similar to unknown proteins
SEQ ID n° 5793	3657	1564	SA-2576.2	1649770-1650459 m	ribosomal protein L1
SEQ ID n° 5794	3658	1566	SA-258.1	303257-303583 p	similar to cell division protein FtsL
SEQ ID n° 5795	3659	1567	SA-2582.2	1123914-1124204 m	similar to unknown protein
SEQ ID n° 5796	3660	1568	SA-2583.3	2056908-2057630 m	similar to unknown proteins
SEQ ID n° 5797	3661	1572	SA-259.1	302295-303242 p	similar to unknown proteins
SEQ ID n° 5798	3662	1575	SA-2596.1	1720081-1720746 m	similar to amino acid ABC transporter (permease)

SEQ ID n° 5799	3663	1576	SA-2597.2	1720770-1721627 m	Unknown, similar to amino acid ABC transporter (binding protein)
SEQ ID n° 5800	3664	1577	SA-2598.3	1077491-1078357 m	similar to phosphate ABC transporter (binding protein)
SEQ ID n° 5801	3665	1578	SA-2599.1	1205070-1206410 m	Similar to drug-export protein
SEQ ID n° 5802	3666	1579	SA-26.1	1051977-1052234 m	Unknown
SEQ ID n° 5803	3667	1580	SA-260.1	300959-302212 p	similar to Gamma-glutamyl phosphate reductase
SEQ ID n° 5804	3668	1589	SA-261.1	300146-300949 p	similar to gamma-glutamyl kinase
SEQ ID n° 5805	3669	1595	SA-263.1	298620-300005 m	similar to Similar to beta-glucoside specific PTS system enzyme
SEQ ID n° 5806	3670	1596	SA-2632.1	1064917-1065144 m	IIBC
SEQ ID n° 5807	3671	1598	SA-264.1	296808-298412 p	Unknown
SEQ ID n° 5808	3672	1602	SA-265.1	296070-296804 p	Transmembrane protein similar to unknown proteins
SEQ ID n° 5809	3673	1604	SA-2651.1	2042706-2042933 p	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5810	3674	1606	SA-2655.1	81684-82052 p	Hypothetical gene
SEQ ID n° 5811	3675	1608	SA-266.1	295745-296050 p	ribosomal protein L14
SEQ ID n° 5812	3676	1613	SA-267.1	293547-295532 p	similar to unknown proteins
SEQ ID n° 5813	3677	1616	SA-2673.1	827405-827671 m	similar to transketolase
SEQ ID n° 5814	3678	1617	SA-268.1	291992-293422 p	Similar to unknown proteins
SEQ ID n° 5815	3679	1622	SA-269.1	290664-292001 p	similar to transcriptional regulatory protein (N-terminal part)
SEQ ID n° 5816	3680	1624	SA-2691.1	2150868-2151116 p	similar to NADH oxidase
SEQ ID n° 5817	3681	1626	SA-270.1	289878-290576 p	similar to phage protein
SEQ ID n° 5818	3682	1627	SA-2703.1	472009-472830 p	similar to glycerol uptake facilitator
SEQ ID n° 5819	3683	1629	SA-2707.1	1510882-1511082 m	similar to unknown protein
SEQ ID n° 5820	3684	1630	SA-2709.1	1511122-1511652 m	50S ribosomal protein L35
SEQ ID n° 5821	3685	1631	SA-271.1	288037-289866 p	translation initiation factor IF-3
SEQ ID n° 5822	3686	1632	SA-2710.1	1511813-1512496 m	glycerol-3-phosphate dehydrogenase
SEQ ID n° 5823	3687	1634	SA-2712.1	1512507-1513031 m	Similar to cytidine monophosphate kinase
SEQ ID n° 5824	3688	1638	SA-272.1	286516-288024 p	Similar to unknown proteins (serine rich)
SEQ ID n° 5825	3689	1641	SA-273.1	286139-286402 p	glycerol kinase
SEQ ID n° 5826	3690	1643	SA-274.1	285793-286050 p	similar to unknown proteins
SEQ ID n° 5827	3691	1646	SA-275.1	283742-285781 p	similar to unknown proteins
SEQ ID n° 5828	3692	1648	SA-2756.1	1714969-1715208 m	glycyl-tRNA synthetase (beta subunit)
SEQ ID n° 5829	3693	1651	SA-2759.1	1715366-1715908 m	Similar to unknown proteins
SEQ ID n° 5830	3694	1652	SA-276.1	283097-283738 p	Similar to unknown proteins
SEQ ID n° 5831	3695	1655	SA-2763.2	1878992-1880002 m	similar to acyl carrier protein phosphodiesterase
SEQ ID n° 5832	3696	1658	SA-2768.2	1349651-1350061 m	glyceraldehyde 3-phosphate dehydrogenase
					similar to unknown proteins

SEQ ID n° 5833	3697	1660	SA-277.1	282183-283097 p	glycyl-tRNA synthetase (alpha subunit)
SEQ ID n° 5834	3698	1663	SA-2772.1	1106781-1107011 m	Unknown
SEQ ID n° 5835	3699	1666	SA-278.2	281328-281870 p	similar to unknown proteins
SEQ ID n° 5836	3700		SA-2787.3	1919967-1920950 m	similar to transcriptional regulator (LacI family)
SEQ ID n° 5837	3701		SA-2789.1	508125-508481 p	Unknown
SEQ ID n° 5838	3702		SA-2790.2	508546-508968 m	Unknown
SEQ ID n° 5839	3703		SA-2791.2	508975-509313 m	similar to transcriptional regulator (phage related)
SEQ ID n° 5840	3704		SA-2799.2	517052-517363 p	Similar to unknown proteins
SEQ ID n° 5841	3705	1668	SA-28.1	730066-732432 p	similar to unknown proteins
SEQ ID n° 5842	3706	1669	SA-280.2	1336040-1339798 m	Similar to pullulanase, Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5843	3707		SA-2801.2	516689-516943 p	Similar to unknown proteins
SEQ ID n° 5844	3708		SA-2803.2	509314-510102 p	hypothetical protein
SEQ ID n° 5845	3709		SA-2804.2	1613541-1614263 m	similar to pseudouridylate synthase
SEQ ID n° 5846	3710		SA-2813.1	236431-236553 p	similar to transcriptional regulator
SEQ ID n° 5847	3711		SA-2815.2	30442-31410 p	similar to ribose-phosphate pyrophosphokinase
SEQ ID n° 5848	3712	1670	SA-282.1	1334988-1335905 m	similar to putative permease
SEQ ID n° 5849	3713		SA-2822.1	2154545-2155606 m	Unknown
SEQ ID n° 5850	3714		SA-2823.2	437597-438925 p	similar to unknown plasmid protein
SEQ ID n° 5851	3715		SA-2829.1	763667-764995 p	Unknown
SEQ ID n° 5852	3716	1671	SA-283.1	1334770-1334946 p	Similar to unknown proteins
SEQ ID n° 5853	3717		SA-2831.1	763222-763533 p	Unknown
SEQ ID n° 5854	3718		SA-2832.1	759654-760991 p	similar to plasmid proteins
SEQ ID n° 5855	3719		SA-2833.1	756757-757404 p	similar to unknown proteins
SEQ ID n° 5856	3720		SA-2834.1	756452-756757 p	Unknown
SEQ ID n° 5857	3721		SA-2837.2	2208347-2209039 m	similar to unknown proteins
SEQ ID n° 5858	3722		SA-2839.1	2111022-2111306 m	chaperonin GroES
SEQ ID n° 5859	3723		SA-2840.1	2057634-2058251 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5860	3724		SA-2842.2	717307-718062 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 5861	3725		SA-2843.1	716495-717295 m	Similar to amino acid ABC transporter (binding protein)
SEQ ID n° 5862	3726		SA-2849.1	125222-125668 p	similar to deoxyuridine 5 -triphosphate nucleotidhydrolase (dUTPase)
SEQ ID n° 5863	3727	1672	SA-285.1	1333769-1334659 m	similar to tRNA isopentenylpyrophosphate transferase
SEQ ID n° 5864	3728		SA-2853.1	88308-88526 p	translation initiation factor IF-1
SEQ ID n° 5865	3729		SA-2854.2	1351575-1351865 p	similar to transposase
SEQ ID n° 5866	3730		SA-2855.1	1363655-1364491 p	similar to transposase, C-terminal part

SEQ ID n° 5867	3731	SA-2856.1	1493128-1493649 m	similar to pyrimidine biosynthetic operon repressor
SEQ ID n° 5868	3732	SA-2857.2	1557873-1558250 p	large conductance mechanosensitive channel protein MscL
SEQ ID n° 5869	3733	SA-286.1	1332440-1333678 m	similar to GTP-binding protein
SEQ ID n° 5870	3734	SA-2860.1	1169149-1169730 m	similar to xanthine phosphoribosyltransferase
SEQ ID n° 5871	3735	SA-2861.1	1245656-1245853 p	similar to hypothetical transcriptional regulator
SEQ ID n° 5872	3736	SA-2862.1	1251589-1251837 m	similar to unknown proteins
SEQ ID n° 5873	3737	SA-2863.1	1251885-1252130 m	similar to unknown proteins
SEQ ID n° 5874	3738	SA-2866.1	762834-763217 p	similar to putative plasmid replication protein
SEQ ID n° 5875	3739	SA-2867.1	762543-762821 p	Unknown
SEQ ID n° 5876	3740	SA-2868.1	761728-762546 p	similar to plasmid partition protein ParA
SEQ ID n° 5877	3741	SA-2869.1	761005-761589 p	Unknown
SEQ ID n° 5878	3742	SA-287.1	1331800-1332447 m	similar to unknown proteins
SEQ ID n° 5879	3743	SA-2870.1	759383-759640 p	Unknown
SEQ ID n° 5880	3744	SA-2871.1	757414-759363 p	similar to plasmid protein
SEQ ID n° 5881	3745	SA-2872.1	755988-756221 p	Unknown
SEQ ID n° 5882	3746	SA-2874.1	753724-755931 p	similar to unknown protein
SEQ ID n° 5883	3747	SA-2875.1	753122-753604 p	Unknown
SEQ ID n° 5884	3748	SA-2877.1	749873-753022 p	similar to plasmid protein
SEQ ID n° 5885	3749	SA-2878.1	749623-749880 p	Unknown
SEQ ID n° 5886	3750	SA-2879.1	748964-749395 p	similar to single-strand binding protein
SEQ ID n° 5887	3751	SA-288.1	1330829-1331758 m	Similar to unknown proteins
SEQ ID n° 5888	3752	SA-2880.1	1036003-1038804 m	similar to surface antigen proteins, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5889	3753	SA-2882.1	745217-745939 p	similar to unknown protein, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5890	3754	SA-2883.1	742879-745140 p	Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5891	3755	SA-2885.1	741782-742864 p	Unknown
SEQ ID n° 5892	3756	SA-2886.1	741424-741771 p	6 No similarity
SEQ ID n° 5893	3757	SA-2887.1	739413-741332 p	similar to ATP-dependent Clp protease ATP-binding subunit ClpA
SEQ ID n° 5894	3758	SA-2888.1	737750-739498 p	similar to DNA topoisomerase III
SEQ ID n° 5895	3759	SA-2890.1	406767-411497 p	similar to plasmid proteins
SEQ ID n° 5896	3760	SA-2891.1	732558-732815 p	Unknown
SEQ ID n° 5897	3761	SA-2893.1	1052360-1054726 m	similar to unknown proteins
SEQ ID n° 5898	3762	SA-2894.1	729554-730012 p	Unknown
SEQ ID n° 5899	3763	SA-2895.1	1055255-1056055 m	similar to unknown proteins

SEQ ID n° 5900	3764	SA-2896.1	728390-728737 p	Unknown
SEQ ID n° 5901	3765	SA-2897.1	728080-728397 p	Unknown
SEQ ID n° 5902	3766	SA-29.1	1054780-1055238 m	Unknown
SEQ ID n° 5903	3767	SA-290.1	1330066-1330827 m	similar to oxidoreductase
SEQ ID n° 5904	3768	SA-2900.1	727411-727686 p	Unknown
SEQ ID n° 5905	3769	SA-2902.1	726446-727288 p	hypothetical start codon
SEQ ID n° 5906	3770	SA-2903.1	725380-726384 p	Unknown
SEQ ID n° 5907	3771	SA-2905.1	724345-725301 m	Unknown
SEQ ID n° 5908	3772	SA-2906.1	724104-724352 p	Unknown
SEQ ID n° 5909	3773	SA-2907.1	723919-724089 p	Unknown
SEQ ID n° 5910	3774	SA-2908.1	723358-723768 p	Unknown
SEQ ID n° 5911	3775	SA-291.1	1327871-1330069 m	similar to single-strand DNA-specific exonuclease
SEQ ID n° 5912	3776	SA-2910.1	722916-723365 p	RecJ
SEQ ID n° 5913	3777	SA-2911.1	1062066-1062512 m	Unknown
SEQ ID n° 5914	3778	SA-2912.1	721852-722280 p	similar to unknown proteins
SEQ ID n° 5915	3779	SA-2913.1	721637-721840 p	Unknown
SEQ ID n° 5916	3780	SA-2914.1	1063411-1064904 m	similar to plasmid replication protein
SEQ ID n° 5917	3781	SA-2916.1	719648-719875 p	Unknown
SEQ ID n° 5918	3782	SA-2917.1	719536-719655 p	Unknown
SEQ ID n° 5919	3783	SA-2918.1	719438-719530 p	hypothetical gene
SEQ ID n° 5920	3784	SA-2919.1	719103-719432 p	Unknown
SEQ ID n° 5921	3785	SA-292.1	1325155-1327719 m	similar to putative hydrolytic protein
SEQ ID n° 5922	3786	SA-2921.3	1558627-1559130 m	similar to unknown proteins
SEQ ID n° 5923	3787	SA-2922.1	1559213-1560019 m	Similar to amino acid ABC transporter (binding protein)
SEQ ID n° 5924	3788	SA-2926.1	1882439-1882909 m	ribosomal protein S7
SEQ ID n° 5925	3789	SA-2928.1	1882931-1883344 m	ribosomal protein S12
SEQ ID n° 5926	3790	SA-293.1	1324514-1325032 m	similar to adenine phosphoribosyltransferase
SEQ ID n° 5927	3791	SA-2939.1	1982190-1982534 m	similar to hypothetical thioredoxin
SEQ ID n° 5928	3792	SA-294.1	1323716-1324396 m	similar to unknown proteins
SEQ ID n° 5929	3793	SA-2940.1	1981772-1982113 m	similar to unknown proteins
SEQ ID n° 5930	3794	SA-2945.1	1765545-1765901 m	similar to unknown proteins
SEQ ID n° 5931	3795	SA-2946.1	1701220-1701810 m	ATP-dependent CLP protease proteolytic subunit
SEQ ID n° 5932	3796	SA-2947.1	1700875-1701120 m	similar to unknown proteins
SEQ ID n° 5933	3797	SA-295.1	1322929-1323612 m	similar to unknown proteins
SEQ ID n° 5934	3798	SA-2951.2	1558347-1558523 m	ribosomal protein S21

SEQ ID n° 5935	3799	SA-2954.1	1350730-1351386 p	similar to unknown proteins
SEQ ID n° 5936	3800	SA-2955.1	1350071-1350568 m	similar to unknown proteins
SEQ ID n° 5937	3801	SA-296.1	1322151-1322939 m	Similar to unknown proteins
SEQ ID n° 5938	3802	SA-2961.1	1136745-1137083 m	similar to N-terminal first 100 amino acids of anaerobic ribonucleotide reductase
SEQ ID n° 5939	3803	SA-297.1	1321039-1322142 m	similar to probable D-amino acid oxidase
SEQ ID n° 5940	3804	SA-2973.2	433313-433570 p	Unknown
SEQ ID n° 5941	3805	SA-2974.2	431344-433293 p	similar to plasmid transfer complex protein
SEQ ID n° 5942	3806	SA-2975.2	427052-427534 p	Unknown
SEQ ID n° 5943	3807	SA-2976.2	423803-426952 p	similar to plasmid proteins
SEQ ID n° 5944	3808	SA-2977.2	419921-422722 p	similar to surface antigen proteins, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5945	3809	SA-2978.1	718598-719092 p	Unknown
SEQ ID n° 5946	3810	SA-2979.1	718373-718561 p	Unknown
SEQ ID n° 5947	3811	SA-298.5	1320126-1320980 m	glucose-1-phosphate thymidyltransferase
SEQ ID n° 5948	3812	SA-2985.2	437152-437463 p	Unknown
SEQ ID n° 5949	3813	SA-2986.2	436764-437147 p	similar to replication initiation protein from Staphylococci plasmid
SEQ ID n° 5950	3814	SA-2987.2	436473-436751 p	Unknown
SEQ ID n° 5951	3815	SA-2988.2	435658-436476 p	similar to plasmid partition protein ParA
SEQ ID n° 5952	3816	SA-2989.1	434935-435519 p	Unknown
SEQ ID n° 5953	3817	SA-299.5	1319519-1320112 m	dTDP-4-keto-6-deoxyglucose-3,5-epimerase
SEQ ID n° 5954	3818	SA-2990.1	433584-434921 p	similar to plasmid proteins
SEQ ID n° 5955	3819	SA-2991.1	430687-431334 p	similar to unknown protein
SEQ ID n° 5956	3820	SA-2992.1	430382-430687 p	Unknown
SEQ ID n° 5957	3821	SA-2993.1	429918-430151 p	Unknown
SEQ ID n° 5958	3822	SA-2995.1	427654-429861 p	similar to unknown proteins
SEQ ID n° 5959	3823	SA-2997.1	423553-423810 p	Unknown
SEQ ID n° 5960	3824	SA-2998.1	422894-423325 p	similar to single-strand binding protein
SEQ ID n° 5961	3825	SA-3.1	1025432-1027381 m	similar to plasmid transfer complex protein
SEQ ID n° 5962	3826	SA-30.1	728737-729537 p	similar to membrane nuclease
SEQ ID n° 5963	3827	SA-3000.1	419147-419869 p	Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5964	3828	SA-3001.1	416809-419070 p	similar to plasmid surface exclusion protein, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 5965	3829	SA-3002.1	415712-416794 p	similar to unknown proteins
SEQ ID n° 5966	3830	SA-3003.1	415354-415701 p	Unknown

SEQ ID n° 5967	3831		SA-3004.1	413430-415262 p	similar to ATP-dependent Clp protease ATP-binding subunit ClpA
SEQ ID n° 5968	3832		SA-3005.1	411680-413428 p	similar to plasmid DNA topoisomerase
SEQ ID n° 5969	3833		SA-3007.2	733632-737567 p	similar to plasmid proteins
SEQ ID n° 5970	3834		SA-3008.1	406491-406748 p	Unknown
SEQ ID n° 5971	3835	1689	SA-301.1	1318266-1319312 m	similar to dTDP-glucose-4,6-dehydratase
SEQ ID n° 5972	3836		SA-3010.1	403999-406365 p	similar to unknown proteins
SEQ ID n° 5973	3837		SA-3011.1	403487-403945 p	Unknown
SEQ ID n° 5974	3838		SA-3012.1	402670-403470 p	similar to unknown proteins
SEQ ID n° 5975	3839		SA-3013.1	402323-402670 p	Unknown
SEQ ID n° 5976	3840		SA-3014.1	402013-402330 p	Unknown
SEQ ID n° 5977	3841		SA-3016.1	401344-401619 p	Unknown
SEQ ID n° 5978	3842		SA-3018.1	400373-401221 p	Unknown
SEQ ID n° 5979	3843		SA-3019.1	399313-400317 p	similar to unknown protein
SEQ ID n° 5980	3844		SA-3021.1	398278-399234 m	Unknown
SEQ ID n° 5981	3845		SA-3022.1	398037-398285 p	Unknown
SEQ ID n° 5982	3846		SA-3023.1	397879-398022 p	Unknown
SEQ ID n° 5983	3847		SA-3025.1	397291-397701 p	Unknown
SEQ ID n° 5984	3848		SA-3027.1	396849-397298 p	Unknown
SEQ ID n° 5985	3849		SA-3028.1	396213-396659 p	similar to unknown proteins
SEQ ID n° 5986	3850		SA-3029.1	395785-396213 p	Unknown
SEQ ID n° 5987	3851		SA-3030.1	395570-395773 p	Unknown
SEQ ID n° 5988	3852		SA-3031.1	393821-395314 p	similar to plasmid replication protein
SEQ ID n° 5989	3853		SA-3033.1	393581-393808 p	Unknown
SEQ ID n° 5990	3854		SA-3034.1	393469-393588 p	Unknown
SEQ ID n° 5991	3855		SA-3035.1	393284-393463 p	Unknown
SEQ ID n° 5992	3856		SA-3036.1	393036-393365 p	Unknown
SEQ ID n° 5993	3857	1691	SA-305.1	1314996-1318214 p	hyaluronate lyase
SEQ ID n° 5994	3858	1692	SA-306.1	1314297-1314779 m	similar to other proteins
SEQ ID n° 5995	3859	1693	SA-307.1	1313846-1314307 m	similar to unknown proteins
SEQ ID n° 5996	3860		SA-3071.1	917842-918033 p	similar to unknown protein
SEQ ID n° 5997	3861		SA-3072.1	733433-733630 p	truncated, C-terminal part
SEQ ID n° 5998	3862		SA-3073.1	732834-733436 p	similar to unknown proteins (N-terminal part)
SEQ ID n° 5999	3863	1694	SA-308.1	1312564-1313754 m	similar to unknown proteins
SEQ ID n° 6000	3864	1695	SA-309.1	1311348-1312574 m	similar to unknown proteins
SEQ ID n° 6001	3865	1696	SA-31.1	1056055-1056402 m	Unknown

SEQ ID n° 6002	3866	1697	SA-311.1	1309556-1311238 m	similar to alpha-acetolactate synthase
SEQ ID n° 6003	3867	1698	SA-312.1	1308823-1309542 m	similar to alpha-acetolactate decarboxylase
SEQ ID n° 6004	3868		SA-3125.1	83021-83206 p	ribosomal protein S14
SEQ ID n° 6005	3869		SA-3126.1	88552-88668 p	ribosomal protein L36
SEQ ID n° 6006	3870	1699	SA-313.2	1307114-1308769 p	similar to hypothetical fibronectin-binding protein
SEQ ID n° 6007	3871	1700	SA-314.2	1305752-1306756 m	similar to unknown proteins
SEQ ID n° 6008	3872	1701	SA-315.1	1304876-1305739 m	Similar to ABC transporter (permease)
SEQ ID n° 6009	3873	1702	SA-316.1	1304115-1304876 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 6010	3874	1703	SA-318.1	1302157-1303818 m	similar to unknown proteins
SEQ ID n° 6011	3875	1704	SA-319.1	1301312-1302100 m	similar to tributyrin esterase
SEQ ID n° 6012	3876	1705	SA-32.1	1056395-1056712 m	Unknown
SEQ ID n° 6013	3877	1706	SA-320.1	1300512-1301222 m	similar to unknown proteins
SEQ ID n° 6014	3878	1708	SA-322.1	1299603-1300274 m	similar to ribose 5-phosphate isomerase
SEQ ID n° 6015	3879	1709	SA-323.1	1298335-1299546 m	similar to phosphopentomutase
SEQ ID n° 6016	3880	1710	SA-324.1	1297877-1298284 m	similar to arsenate reductase (hypothetical)
SEQ ID n° 6017	3881	1711	SA-325.1	1297029-1297838 m	similar to purine nucleoside phosphorylase
SEQ ID n° 6018	3882	1712	SA-326.1	1295771-1297027 m	putative transport protein
SEQ ID n° 6019	3883	1713	SA-327.1	1295077-1295787 m	similar to purine-nucleoside phosphorylase
SEQ ID n° 6020	3884	1714	SA-328.1	1294301-1295068 m	similar to unknown proteins
SEQ ID n° 6021	3885	1715	SA-329.2	1293353-1294276 p	similar to transcriptional regulator (LysR family)
SEQ ID n° 6022	3886	1716	SA-330.2	60992-61249 p	similar to unknown proteins
SEQ ID n° 6023	3887	1717	SA-331.2	59909-60982 p	similar to phosphoribosylaminoimidazole carboxylase II
SEQ ID n° 6024	3888	1718	SA-333.1	59431-59922 p	similar to phosphoribosylaminoimidazole carboxylase
SEQ ID n° 6025	3889	1719	SA-334.1	57888-59150 p	similar to phosphoribosylamine-glycine ligase
SEQ ID n° 6026	3890	1720	SA-335.1	56801-57604 m	similar to unknown proteins
SEQ ID n° 6027	3891	1721	SA-336.1	55801-56778 p	similar to acetyl xylan esterase (hypothetical)
SEQ ID n° 6028	3892	1722	SA-337.1		
SEQ ID n° 6029	3893	1723	SA-338.1	54912-55793 p	similar to glucose kinase and to XylR transcriptional regulator
SEQ ID n° 6030	3894	1724	SA-339.1	53978-54895 p	similar to N-acetylneuraminate lyase
SEQ ID n° 6031	3895	1725	SA-34.1	53319-53981 p	similar to unknown proteins
SEQ ID n° 6032	3896	1726	SA-340.1	1057106-1057381 m	Unknown
SEQ ID n° 6033	3897	1727	SA-342.1	52856-53299 p	similar to unknown proteins
SEQ ID n° 6034	3898	1728	SA-343.1	52013-52843 p	Similar to ABC transporter (permease)
SEQ ID n° 6035	3899	1729	SA-344.1	51116-52003 p	Similar to sugar ABC transporter (permease)
SEQ ID n° 6036	3900	1730	SA-345.1	49712-51028 p	Similar to ABC transporter (binding protein)
				48967-49665 p	similar to unknown proteins

SEQ ID n° 6037	3901	1731	SA-347.1	47416-48720 p	group B streptococcal surface immunogenic protein
SEQ ID n° 6038	3902	1732	SA-348.1	46370-47269 p	similar to Streptococcus equi zooxin A endopeptidase
SEQ ID n° 6039	3903	1733	SA-351.1	44630-46177 p	similar to bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase
SEQ ID n° 6040	3904	1734	SA-352.1	43858-44610 p	similar to unknown proteins
SEQ ID n° 6041	3905	1735	SA-353.1	43287-43835 p	similar to Phosphoribosylglycinamide formyltransferase
SEQ ID n° 6042	3906	1736	SA-356.1	42097-43119 p	similar to phosphoribosylformylglycinamide cyclo-ligase
SEQ ID n° 6043	3907	1737	SA-357.1	40615-42069 p	similar to phosphoribosylpyrophosphate amidotransferase
SEQ ID n° 6044	3908	1738	SA-36.1	1057504-1058352 m	Unknown
SEQ ID n° 6045	3909	1740	SA-363.1	36656-40381 p	similar to Phosphoribosylformylglycinamide synthase
SEQ ID n° 6046	3910	1741	SA-364.2	35829-36533 p	similar to phosphoribosylaminoimidazole succinocarboxamide synthase
SEQ ID n° 6047	3911	1742	SA-366.2	562906-565167 m	similar to ATP-dependent protease ClpE
SEQ ID n° 6048	3912	1743	SA-367.1	562265-562720 m	Similar to unknown proteins
SEQ ID n° 6049	3913	1744	SA-368.1	561899-562201 m	similar to unknown proteins
SEQ ID n° 6050	3914	1745	SA-37.1	1058408-1059412 m	similar to unknown protein
SEQ ID n° 6051	3915	1746	SA-370.1	558990-561782 p	isoleucyl-tRNA synthetase
SEQ ID n° 6052	3916	1747	SA-371.1	557935-558705 p	similar to cell division protein DivIVA
SEQ ID n° 6053	3917	1748	SA-372.1	557137-557925 p	similar to unknown proteins
SEQ ID n° 6054	3918	1749	SA-373.1	556881-557135 p	Similar to unknown proteins
SEQ ID n° 6055	3919	1750	SA-374.1	556273-556878 p	Similar to unknown proteins
SEQ ID n° 6056	3920	1751	SA-375.1	555590-556261 p	Similar to unknown proteins
SEQ ID n° 6057	3921	1752	SA-376.1	554301-555581 p	cell division protein FtsZ
SEQ ID n° 6058	3922	1753	SA-377.1	552906-554279 p	Similar to cell division protein FtsA
SEQ ID n° 6059	3923	1754	SA-379.1	551497-552633 p	Similar to cell division protein DivIB
SEQ ID n° 6060	3924	1755	SA-380.1	550417-551493 p	Similar to undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase
SEQ ID n° 6061	3925	1756	SA-381.1	549059-550414 p	Similar to UDP-N-acetylmuramoylalanine--D-glutamate ligase
SEQ ID n° 6062	3926	1757	SA-382.1	548684-548929 p	Similar to unknown proteins
SEQ ID n° 6063	3927	1758	SA-384.1	546798-548639 p	similar to putative GTP-binding elongation factor
SEQ ID n° 6064	3928	1759	SA-385.1	546186-546566 p	similar to unknown proteins
SEQ ID n° 6065	3929	1760	SA-387.1	545206-546174 p	similar to glucose kinase
SEQ ID n° 6066	3930	1761	SA-388.1	545000-545209 p	similar to unknown proteins

SEQ ID n° 6067	3931	1762	SA-389.1	544450-544845 m	similar to unknown proteins
SEQ ID n° 6068	3932	1763	SA-39.1	1059491-1060447 p	Unknown
SEQ ID n° 6069	3933	1764	SA-390.1	543817-544449 p	similar to endonuclease III
SEQ ID n° 6070	3934	1765	SA-391.1	542475-543704 p	similar to other proteins
SEQ ID n° 6071	3935	1766	SA-392.1	541439-542482 p	Similar to Acetyl-CoA acetyltransferase (truncated)
SEQ ID n° 6072	3936	1767	SA-393.1	540863-541378 p	Unknown
SEQ ID n° 6073	3937	1768	SA-394.1	539746-540738 m	similar to biotin synthetase
SEQ ID n° 6074	3938	1769	SA-395.1	539206-539745 m	similar to other proteins
SEQ ID n° 6075	3939	1770	SA-396.1	538572-539138 p	similar to Anthranilate synthase component II (Glutamine amido-transferase)
SEQ ID n° 6076	3940	1771	SA-397.1	536705-538444 p	similar to ABC transporter (exporter) (ATP-binding protein)
SEQ ID n° 6077	3941	1772	SA-398.1	534970-536715 p	similar to ABC transporter (exporter) (ATP-binding protein)
SEQ ID n° 6078	3942	1773	SA-399.1	534259-534774 p	Similar to unknown proteins
SEQ ID n° 6079	3943	1774	SA-4.2	1027391-1028038 m	similar to unknown protein
SEQ ID n° 6080	3944	1775	SA-40.1	1060440-1060688 m	Unknown
SEQ ID n° 6081	3945	1776	SA-400.1	533153-534259 p	similar to unknown proteins
SEQ ID n° 6082	3946	1777	SA-401.2	532524-533102 p	similar to unknown proteins
SEQ ID n° 6083	3947	1778	SA-402.2	531121-532455 p	Similar to unknown proteins
SEQ ID n° 6084	3948	1779	SA-403.1	529973-531046 p	similar to unknown proteins
SEQ ID n° 6085	3949	1780	SA-405.1	529498-529983 p	similar to phosphopantetheine adenylyltransferase
SEQ ID n° 6086	3950	1781	SA-407.1	529211-529501 p	Unknown
SEQ ID n° 6087	3951	1782	SA-409.1	528447-529199 p	similar to unknown proteins
SEQ ID n° 6088	3952	1783	SA-41.1	1060703-1060846 m	Unknown
SEQ ID n° 6089	3953	1784	SA-410.1	528080-528529 m	Similar to unknown proteins
SEQ ID n° 6090	3954	1785	SA-411.2	527043-528035 p	Similar to asparagine synthetase
SEQ ID n° 6091	3955	1786	SA-412.2	1663493-1664140 p	similar to metal-dependent transcriptional regulator
SEQ ID n° 6092	3956	1787	SA-413.1	1664180-1664869 m	Similar to 5 -methylthioadenosine nucleosidase/S-adenosylhomocysteine nucleosidase
SEQ ID n° 6093	3957	1788	SA-414.1	1664879-1665148 m	Similar to unknown proteins
SEQ ID n° 6094	3958	1789	SA-415.1	1665148-1665702 m	Similar to ADP-ribose pyrophosphatase
SEQ ID n° 6095	3959	1790	SA-416.1	1665723-1667102 m	similar to UDP-N-acetylglucosamine pyrophosphorylase
SEQ ID n° 6096	3960	1791	SA-417.1	1667346-1667759 m	Similar to unknown proteins
SEQ ID n° 6097	3961	1792	SA-418.1	1667772-1668149 m	Similar to unknown proteins
SEQ ID n° 6098	3962	1793	SA-419.1	1668238-1669194 m	similar to oxidoreductase
SEQ ID n° 6099	3963	1794	SA-42.1	1061024-1061434 m	Unknown
SEQ ID n° 6100	3964	1795	SA-420.1	1669191-1669445 m	similar to unknown proteins C-terminal end

SEQ ID n° 6101	3965	1797	SA-422.1	1669406-1669648 m	Unknown
SEQ ID n° 6102	3966	1799	SA-424.1	1669887-1670585 m	similar to oxidoreductases
SEQ ID n° 6103	3967	1800	SA-425.1	1670578-1670814 m	Similar to unknown proteins
SEQ ID n° 6104	3968	1801	SA-426.1	1670904-1671152 m	Similar to unknown proteins
SEQ ID n° 6105	3969	1802	SA-427.1	1671350-1671850 m	similar to other proteins
SEQ ID n° 6106	3970	1803	SA-428.1	1672090-1672545 m	similar to unknown proteins (C-terminal end)
SEQ ID n° 6107	3971	1804	SA-429.1	1672546-1673502 m	Similar to other proteins (N-terminal part)
SEQ ID n° 6108	3972	1805	SA-430.1	1673642-1675801 m	Similar to unknown proteins
SEQ ID n° 6109	3973	1806	SA-431.1	1675808-1677253 m	Similar to other proteins
SEQ ID n° 6110	3974	1807	SA-432.1	1677246-1677923 m	Unknown
SEQ ID n° 6111	3975	1808	SA-433.1	1677975-1678586 m	Unknown
SEQ ID n° 6112	3976	1809	SA-434.2	1679034-1680371 m	Similar to branched-chain amino acid transporter
SEQ ID n° 6113	3977	1810	SA-435.1	1680567-1682564 m	methionyl-tRNA synthetase
SEQ ID n° 6114	3978	1811	SA-436.1	1682707-1683582 p	Similar to other proteins
SEQ ID n° 6115	3979	1812	SA-437.1	1683740-1684435 m	Similar to unknown proteins
SEQ ID n° 6116	3980	1813	SA-438.1	1684771-1685988 p	Similar to PTS enzyme IIC
SEQ ID n° 6117	3981	1814	SA-439.1	1685972-1686814 p	Similar to unknown proteins
SEQ ID n° 6118	3982	1815	SA-44.1	1061427-1061876 m	Unknown
SEQ ID n° 6119	3983	1816	SA-441.1	1686891-1687718 p	Similar to 3'-exo-deoxyribonuclease
SEQ ID n° 6120	3984	1817	SA-442.1	1687759-1688115 m	Similar to other proteins
SEQ ID n° 6121	3985	1818	SA-443.1	1688117-1688593 m	similar to O6-alkylguanine-DNA-alkyltransferase
SEQ ID n° 6122	3986	1819	SA-445.1	1688649-1689830 m	similar to phosphoglycerate dehydrogenase
SEQ ID n° 6123	3987	1820	SA-446.1	1689892-1690440 m	similar to unknown proteins
SEQ ID n° 6124	3988	1821	SA-447.1	1690509-1691600 m	similar to phosphoserine aminotransferase
SEQ ID n° 6125	3989	1822	SA-448.1	1691733-1692368 p	Similar to other proteins
SEQ ID n° 6126	3990	1823	SA-45.1	72280-722726 p	Unknown
SEQ ID n° 6127	3991	1824	SA-450.1	1692638-1693501 m	Similar to unknown proteins
SEQ ID n° 6128	3992	1825	SA-451.1	1693507-1693833 m	Similar to unknown proteins
SEQ ID n° 6129	3993	1826	SA-452.1	1693864-1694727 m	similar to DNA polymerase III (delta subunit)
SEQ ID n° 6130	3994	1827	SA-453.1	1694747-1695382 m	similar to thymidylate kinase
SEQ ID n° 6131	3995	1828	SA-454.1	1695471-1696130 m	Similar to acetoin dehydrogenase
SEQ ID n° 6132	3996	1829	SA-455.1	1696149-1696859 m	similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 6133	3997	1830	SA-456.1	1696859-1697623 m	similar to amino acid ABC transporter (ATP-binding protein)
SEQ ID n° 6134	3998	1831	SA-457.1	1697624-1698577 m	similar to ABC transporter (permease)
SEQ ID n° 6135	3999	1832	SA-458.2	1698580-1699449 m	similar to amino acid ABC transporter (permease)
SEQ ID n° 6136	4000	1833	SA-46.1	1062512-1062940 m	Unknown

SEQ ID n° 6137	4001	1834	SA-460.3	1699555-1700721 m	similar to branched-chain amino acid ABC transporter, amino acid-binding protein
SEQ ID n° 6138	4002	1835	SA-462.1	365287-365721 p	similar to transcriptional regulator (MarR family)
SEQ ID n° 6139	4003	1836	SA-463.1	365721-366692 p	similar to beta-ketoacyl-ACP synthase III
SEQ ID n° 6140	4004	1837	SA-464.1	366750-366974 p	similar to acyl carrier protein
SEQ ID n° 6141	4005	1838	SA-465.1	367129-368088 p	similar to putative trans-2-enoyl-ACP reductase II
SEQ ID n° 6142	4006	1839	SA-466.1	368108-369034 p	similar to malonyl CoA-acyl carrier protein transacylase
SEQ ID n° 6143	4007	1840	SA-467.1	369043-369777 p	similar to beta-ketoacyl-ACP reductase
SEQ ID n° 6144	4008	1841	SA-468.1	369793-371025 p	similar to 3-oxoacyl-acyl-carrier protein synthase
SEQ ID n° 6145	4009	1842	SA-469.1	371027-371527 p	similar to biotin carboxyl carrier protein
SEQ ID n° 6146	4010	1843	SA-47.1	719888-721381 p	similar to plasmid replication protein E
SEQ ID n° 6147	4011	1844	SA-470.1	371524-371946 p	similar to beta-hydroxyacyl-ACP dehydratase
SEQ ID n° 6148	4012	1845	SA-471.1	371984-373354 p	similar to acetyl-CoA carboxylase biotin carboxylase subunit
SEQ ID n° 6149	4013	1846	SA-472.1	373363-374238 p	similar to acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
SEQ ID n° 6150	4014	1847	SA-473.1	374231-375004 p	similar to acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
SEQ ID n° 6151	4015	1848	SA-474.1	375474-376106 p	similar to sakacin A production response regulator [Streptococcus mutans] hypothetical
SEQ ID n° 6152	4016	1849	SA-475.1	376152-377429 m	seryl-tRNA synthetase
SEQ ID n° 6153	4017	1850	SA-477.1	377720-378787 p	similar to unknown protein
SEQ ID n° 6154	4018	1851	SA-478.1	378825-379187 m	similar to unknown protein
SEQ ID n° 6155	4019	1853	SA-480.1	379306-380217 m	similar to mannose-specific PTS enzyme IID
SEQ ID n° 6156	4020	1854	SA-481.1	380232-381044 m	similar to mannose-specific PTS enzyme IIC
SEQ ID n° 6157	4021	1855	SA-483.1	381077-382087 m	similar to mannose-specific PTS enzyme IID
SEQ ID n° 6158	4022	1856	SA-484.1	382390-383202 m	similar to unknown protein
SEQ ID n° 6159	4023	1857	SA-485.1	383291-383875 p	putative transmembrane protein
SEQ ID n° 6160	4024	1858	SA-486.1	383965-384576 p	similar to unknown transmembrane protein
SEQ ID n° 6161	4025	1860	SA-488.1	384679-386100 p	similar to unknown transmembrane protein
SEQ ID n° 6162	4026	1861	SA-489.1	386249-386692 p	similar to unknown protein
SEQ ID n° 6163	4027	1862	SA-49.2	1065262-1065354 m	Unknown
SEQ ID n° 6164	4028	1863	SA-490.1	386685-387206 p	similar to unknown protein
SEQ ID n° 6165	4029	1864	SA-491.1	387215-388522 p	similar to transcription regulator, hypothetical.
SEQ ID n° 6166	4030	1865	SA-493.1	388586-388882 m	similar to unknown protein
SEQ ID n° 6167	4031	1866	SA-494.1	388879-389298 m	similar to cell-cycle regulation histidine triad (HIT) protein
SEQ ID n° 6168	4032	1867	SA-495.1	389634-390137 p	Unknown

SEQ ID n° 6169	4033	1868	SA-497.2	1252542-1254821 m	similar to ATP-dependent DNA helicase
SEQ ID n° 6170	4034	1869	SA-498.1	1254927-1255313 m	similar to unknown proteins
SEQ ID n° 6171	4035	1870	SA-499.1	1255446-1256771 m	similar to uracil permease
SEQ ID n° 6172	4036	1871	SA-5.2	1028038-1028343 m	Unknown
SEQ ID n° 6173	4037	1872	SA-50.1	1065137-1065256 m	Unknown
SEQ ID n° 6174	4038	1873	SA-500.1	1257368-1258714 p	similar to probable amino-acid transporter
SEQ ID n° 6175	4039	1874	SA-501.1	1258778-1260013 p	similar to putative cation efflux system protein
SEQ ID n° 6176	4040	1875	SA-502.1	1260172-1260564 p	similar to unknown proteins
SEQ ID n° 6177	4041	1876	SA-503.2	1260545-1261240 p	similar to unknown proteins
SEQ ID n° 6178	4042	1877	SA-504.2	1261309-1261932 p	similar to unknown proteins
SEQ ID n° 6179	4043	1878	SA-505.1	1262334-1262480 m	Unknown
SEQ ID n° 6180	4044	1879	SA-506.1	1262533-1262964 p	Unknown
SEQ ID n° 6181	4045	1880	SA-507.1	1263261-1263695 p	Unknown
SEQ ID n° 6182	4046	1881	SA-508.2	1264088-1264876 p	similar to repressor protein - phage associated
SEQ ID n° 6183	4047	1882	SA-509.2	1211420-1212304 m	Similar to putative pseudouridine synthase
SEQ ID n° 6184	4048	1883	SA-51.1	1065360-1065689 m	Unknown
SEQ ID n° 6185	4049	1884	SA-510.1	1212301-1213137 m	Similar to unknown protein
SEQ ID n° 6186	4050	1885	SA-511.1	1213112-1213783 m	Similar to unknown protein
SEQ ID n° 6187	4051	1886	SA-512.1	1213893-1214465 p	Similar to unknown protein
SEQ ID n° 6188	4052	1887	SA-513.1	1214842-1215616 p	Similar to Phosphoribosylpyrophosphate synthetase
SEQ ID n° 6189	4053	1890	SA-517.1	362167-363519 m	similar to aspartokinase
SEQ ID n° 6190	4054	1891	SA-518.1	363613-364263 p	similar to unknown protein
SEQ ID n° 6191	4055	1892	SA-519.3	364400-365191 p	similar to enoyl-CoA isomerase
SEQ ID n° 6192	4056	1893	SA-52.1	392531-393025 p	Unknown
SEQ ID n° 6193	4057	1894	SA-520.2	507494-507787 p	Similar to unknown proteins
SEQ ID n° 6194	4058	1895	SA-523.4	503448-506828 p	similar to alpha protein, putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 6195	4059	1896	SA-524.1	502012-503205 m	similar to transcriptional regulator (AraC/XylS family)
SEQ ID n° 6196	4060	1897	SA-526.1	501293-501817 p	similar to transcriptional regulator tetR-family
SEQ ID n° 6197	4061	1898	SA-527.3	500298-501161 m	similar to cation efflux system protein
SEQ ID n° 6198	4062	1899	SA-528.3	499830-500213 m	similar to similar to oxydoreductases, N-terminal part
SEQ ID n° 6199	4063	1900	SA-529.1	499359-499829 m	similar to oxydoreductase (C-terminal part)
SEQ ID n° 6200	4064	1901	SA-53.1	392306-392494 p	Unknown
SEQ ID n° 6201	4065	1902	SA-530.1	498794-499249 m	similar to alcohol dehydrogenase (N-terminal part)
SEQ ID n° 6202	4066	1903	SA-531.1	498213-498908 m	similar to alcohol dehydrogenase (C-terminal part)
SEQ ID n° 6203	4067	1904	SA-532.1	497808-498197 m	similar to transcriptional regulator (MerR family)

SEQ ID n° 6204	4068	1905	SA-533.1	497403-497798 m	Similar to other proteins
SEQ ID n° 6205	4069	1906	SA-534.1	497062-497379 m	Similar to decarboxylase
SEQ ID n° 6206	4070	1907	SA-535.1	496691-496975 m	Similar to unknown proteins
SEQ ID n° 6207	4071	1908	SA-536.1	495961-496452 m	Similar to hypothetical transcriptional regulators
SEQ ID n° 6208	4072	1909	SA-537.1	495528-495917 p	similar to unknown proteins
SEQ ID n° 6209	4073	1910	SA-538.1	492348-495515 p	Similar to surface proteins, putative peptidoglycan bound protein (LPXTS motif)
SEQ ID n° 6210	4074	1911	SA-54.2	391173-392213 p	similar to ABC transporter (permease)
SEQ ID n° 6211	4075	1912	SA-540.1	490117-492285 p	Similar to ribonucleoside-diphosphate reductase 2 alpha subunit
SEQ ID n° 6212	4076	1913	SA-541.1	489702-490115 p	Similar to unknown proteins
SEQ ID n° 6213	4077	1914	SA-542.2	488691-489701 p	similar to ribonucleoside-diphosphate reductase beta chain
SEQ ID n° 6214	4078	1915	SA-544.1	487230-488138 p	similar to rhamnosyltransferase
SEQ ID n° 6215	4079	1918	SA-547.2	483319-487020 p	Putative peptidoglycan bound protein (LPXTS motif) similar to C5A peptidase
SEQ ID n° 6216	4080	1919	SA-548.2	482703-483131 p	similar to unknown proteins
SEQ ID n° 6217	4081	1920	SA-549.1	481905-482669 p	Similar to purine nucleoside phosphorylase
SEQ ID n° 6218	4082	1921	SA-55.2	390446-391171 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 6219	4083	1922	SA-550.1	480283-481638 p	Similar to other proteins including hypothetical methyltransferases
SEQ ID n° 6220	4084	1923	SA-551.1	479408-480184 m	Similar to unknown proteins
SEQ ID n° 6221	4085	1924	SA-552.1	478791-479324 m	Similar to unknown proteins
SEQ ID n° 6222	4086	1925	SA-554.2	1577255-1578580 m	Similar to unknown proteins
SEQ ID n° 6223	4087	1926	SA-555.1	1578573-1580081 m	Similar to putative glucosyl transferase
SEQ ID n° 6224	4088	1927	SA-557.1	1580095-1582482 m	Similar to preprotein translocase secA
SEQ ID n° 6225	4089	1928	SA-558.1	1582469-1583461 m	Similar to unknown proteins
SEQ ID n° 6226	4090	1929	SA-559.1	1583458-1585017 m	Similar to unknown protein
SEQ ID n° 6227	4091	1931	SA-560.1	1585024-1586568 m	Similar to unknown protein
SEQ ID n° 6228	4092	1932	SA-561.1	1586568-1587797 m	Similar to preprotein translocase secY
SEQ ID n° 6229	4093	1933	SA-562.1	1587921-1589117 m	Similar to hypothetical glycosyl transferase
SEQ ID n° 6230	4094	1934	SA-563.1	1589178-1590053 m	Similar to putative glycosyltransferase
SEQ ID n° 6231	4095	1935	SA-565.1	1590046-1591230 m	Similar to putative glycosyl transferase
SEQ ID n° 6232	4096	1936	SA-566.1	1591220-1592461 m	Similar to putative glycosyl transferase
SEQ ID n° 6233	4097	1937	SA-567.1	1592458-1593663 m	Similar to putative glycosyl transferase
SEQ ID n° 6234	4098	1938	SA-568.1	1593672-1594679 m	Similar to unknown proteins
SEQ ID n° 6235	4099	1940	SA-57.1	2155624-2156007 m	Unknown

SEQ ID n° 6236	4100	1942	SA-571.1	1595010-1598942 m	Similar to streptococcal hemagglutinin from <i>Streptococcus gordonii</i> , Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 6237	4101	1943	SA-572.1	159325-1600821 p	Similar to transcription regulator RofA related
SEQ ID n° 6238	4102	1944	SA-573.1	1600918-1602909 m	exinuclease ABC chain B
SEQ ID n° 6239	4103	1945	SA-574.1	1602970-1603890 m	Similar to unknown proteins
SEQ ID n° 6240	4104	1946	SA-575.1	1604073-1606256 p	glutamine ABC transporter permease and substrate binding protein
SEQ ID n° 6241	4105	1947	SA-576.1	1606256-1606996 p	glutamine ABC transporter ATP-binding protein
SEQ ID n° 6242	4106	1948	SA-577.1	1607145-1607495 p	Unknown
SEQ ID n° 6243	4107	1949	SA-578.1	1607587-1607745 m	Similar to unknown proteins
SEQ ID n° 6244	4108	1950	SA-579.1	1607771-1609084 m	Similar to GTP-binding protein
SEQ ID n° 6245	4109	1952	SA-582.1	1609451-1610692 p	Similar to aminopeptidase
SEQ ID n° 6246	4110	1954	SA-584.1	1610726-1611304 m	Putative peptidoglycan bound serine rich protein (LPXTG motif)
SEQ ID n° 6247	4111	1955	SA-585.4	1611414-1613456 m	similar to amidase or hydrolase, putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 6248	4112	1957	SA-589.2	1797911-1798792 m	Similar to fructokinase
SEQ ID n° 6249	4113	1959	SA-590.1	1796846-1797793 m	Similar to mannose-6-phosphate isomerase
SEQ ID n° 6250	4114	1960	SA-591.1	1794209-1796737 m	Similar to preprotein translocase SecA subunit
SEQ ID n° 6251	4115	1961	SA-592.2	1793076-1794083 m	similar to 2-dehydro-3-deoxyphosphoheptone aldolase
SEQ ID n° 6252	4116	1962	SA-593.2	1792692-1793051 m	similar to holo-acyl-carrier protein synthase
SEQ ID n° 6253	4117	1963	SA-594.2	1791595-1792695 m	similar to alanine racemase
SEQ ID n° 6254	4118	1964	SA-597.1	1789964-1791502 m	Similar to immunogenic secreted protein
SEQ ID n° 6255	4119	1965	SA-598.1	1787869-1789884 m	Similar to ATP-dependent DNA helicase RecG
SEQ ID n° 6256	4120	1966	SA-599.1	1786664-1787578 m	Similar to oxidoreductase
SEQ ID n° 6257	4121	1967	SA-6.1	1028574-1028807 m	Unknown
SEQ ID n° 6258	4122	1968	SA-60.1	2156179-2156709 m	hypothetical gene
SEQ ID n° 6259	4123	1969	SA-600.1	1785688-1786566 m	Similar to shikimate 5-dehydrogenase
SEQ ID n° 6260	4124	1970	SA-601.1	1784691-1785653 p	Similar to L-asparaginase
SEQ ID n° 6261	4125	1971	SA-602.1	1783240-1784622 m	Similar to unknown proteins
SEQ ID n° 6262	4126	1972	SA-603.2	1782732-1783184 p	Similar to unknown proteins
SEQ ID n° 6263	4127	1973	SA-604.2	1781253-1782464 m	Similar to putative aminotransferase
SEQ ID n° 6264	4128	1974	SA-605.1	1780342-1781127 m	Similar to transcriptional regulator (CodY family)
SEQ ID n° 6265	4129	1975	SA-606.1	1779727-1780275 m	Similar to other proteins
SEQ ID n° 6266	4130	1976	SA-607.1	1778715-1779680 p	Similar to 3-hydroxyacyl-CoA dehydrogenase
SEQ ID n° 6267	4131	1977	SA-608.1	1777897-1778526 m	Similar to unknown proteins

SEQ ID n° 6268	4132	1978	SA-609.1	177056-1777886 m	Similar to unknown proteins
SEQ ID n° 6269	4133	1979	SA-611.1	1774398-1777043 m	Similar to Pyruvate Phosphate Dikinase
SEQ ID n° 6270	4134	1980	SA-612.1	1773958-1774260 m	Similar to Glu-tRNA Gln amidotransferase subunit C
SEQ ID n° 6271	4135	1981	SA-613.1	1772492-1773958 m	Similar to Glutaryl-tRNA Gln amidotransferase subunit A
SEQ ID n° 6272	4136	1982	SA-614.1	1771050-1772492 m	Similar to Glu-tRNA amidotransferase subunit B
SEQ ID n° 6273	4137	1983	SA-615.1	1770006-1770920 m	Similar to unknown proteins
SEQ ID n° 6274	4138	1984	SA-616.1	1769362-1769922 m	Similar to unknown proteins
SEQ ID n° 6275	4139	1985	SA-617.1	1768244-1769362 m	Similar to unknown proteins
SEQ ID n° 6276	4140	1986	SA-618.1	1767834-1768151 m	Similar to unknown proteins
SEQ ID n° 6277	4141	1987	SA-62.1	2156559-2157413 p	similar to integrase, C-terminal part
SEQ ID n° 6278	4142	1988	SA-620.1	1767072-1767704 m	Similar to probable nicotinate-nucleotide adenylyltransferase
SEQ ID n° 6279	4143	1989	SA-621.1	1766488-1767075 m	Similar to unknown proteins
SEQ ID n° 6280	4144	1990	SA-622.2	1765903-1766424 m	similar to unknown proteins
SEQ ID n° 6281	4145	1991	SA-623.2	818453-819139 p	Similar to unknown proteins
SEQ ID n° 6282	4146	1992	SA-624.1	817104-818324 p	Similar to transporter (antiporter)
SEQ ID n° 6283	4147	1993	SA-625.1	815550-816917 p	Similar to UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase
SEQ ID n° 6284	4148	1994	SA-626.1	814357-815403 p	Similar to D-alanine-D-alanine ligase
SEQ ID n° 6285	4149	1995	SA-627.1	813620-814216 p	Similar to recombination protein RecR
SEQ ID n° 6286	4150	1996	SA-628.1	811563-813605 p	Similar to penicillin binding protein 2B
SEQ ID n° 6287	4151	1997	SA-629.1	810739-811431 p	Similar to phosphoglycerate mutase
SEQ ID n° 6288	4152	1998	SA-63.1	2157770-2158441 p	Unknown
SEQ ID n° 6289	4153	1999	SA-630.1	809804-810562 p	Similar to triosephosphate isomerase
SEQ ID n° 6290	4154	2001	SA-633.1	808427-809623 p	elongation factor Tu
SEQ ID n° 6291	4155	2002	SA-635.1	806807-808075 p	Similar to cell division protein FtsW and to RodA protein
SEQ ID n° 6292	4156	2003	SA-636.1	803906-806701 p	Similar to phosphoenolpyruvate carboxylase
SEQ ID n° 6293	4157	2004	SA-637.1	801898-803697 m	Similar to oligopeptidase
SEQ ID n° 6294	4158	2005	SA-638.1	801450-801839 p	Similar to unknown proteins
SEQ ID n° 6295	4159	2006	SA-639.1	800942-801466 p	Similar to unknown proteins
SEQ ID n° 6296	4160	2007	SA-64.1	2158526-2159197 p	Similar to two-component response regulator
SEQ ID n° 6297	4161	2008	SA-640.1	799937-800785 m	Similar to bacteriophage endolysin
SEQ ID n° 6298	4162	2009	SA-641.1	799287-799904 p	Similar to other proteins
SEQ ID n° 6299	4163	2010	SA-642.1	798519-798992 m	Similar to transcriptional regulator
SEQ ID n° 6300	4164	2011	SA-643.1	797856-798497 m	Similar to putative phosphoglycerate mutase
SEQ ID n° 6301	4165	2012	SA-644.1	796918-797820 p	Similar to unknown proteins
SEQ ID n° 6302	4166	2013	SA-646.1	795253-796743 m	lysyl-tRNA synthetase

SEQ ID n° 6303	4167	2014	SA-647.1	794708-795178 p	Similar to riboflavin synthase complex beta chain
SEQ ID n° 6304	4168	2015	SA-648.1	793500-794693 p	Similar to GTP cyclohydrazase II/ 3,4-dihydroxy-2-butanone-4-phosphate synthase
SEQ ID n° 6305	4169	2016	SA-649.1	792832-793482 p	Similar to riboflavin synthase alpha chain
SEQ ID n° 6306	4170	2017	SA-651.1	2159181-2160545 p	Similar to two-component sensor histidine kinase
SEQ ID n° 6307	4171	2018	SA-651.1	791742-792851 p	Similar to riboflavin specific deaminase (diamino-6-(5-phosphoribosylamino)uracil reductase)
SEQ ID n° 6308	4172	2019	SA-653.1	789673-791013 m	Similar to manganese transporter
SEQ ID n° 6309	4173	2020	SA-654.2	788774-789571 p	Similar to unknown proteins
SEQ ID n° 6310	4174	2021	SA-655.2	788438-788650 p	Similar to unknown proteins
SEQ ID n° 6311	4175	2022	SA-656.2	787025-788311 p	Similar to putative peptidases
SEQ ID n° 6312	4176	2023	SA-657.2	785969-786895 p	Similar to putative proteases
SEQ ID n° 6313	4177	2024	SA-658.2	2026535-2028718 m	Similar to glucose-specific PTS enzyme IIA/ABC
SEQ ID n° 6314	4178	2027	SA-660.1	2025666-2026481 m	Similar to unknown protein
SEQ ID n° 6315	4179	2028	SA-661.1	2024637-2025389 m	Similar to two-component response regulator
SEQ ID n° 6316	4180	2029	SA-662.1	2024269-2024538 m	Similar to two-component sensor histidine kinase (C-terminal part)
SEQ ID n° 6317	4181	2030	SA-663.1	2023298-2024146 m	Similar to ABC transporter (ATP-binding protein)
SEQ ID n° 6318	4182	2031	SA-664.1	2022138-2023292 m	Putative transmembrane protein
SEQ ID n° 6319	4183	2033	SA-666.1	2020875-2021927 m	Similar to other proteins, putative transmembrane protein
SEQ ID n° 6320	4184	2034	SA-667.1	2020203-2020628 m	Similar to mannose-specific PTS enzyme IIA
SEQ ID n° 6321	4185	2036	SA-669.1	2019694-2020185 m	Similar to mannose-specific PTS enzyme IIB
SEQ ID n° 6322	4186	2037	SA-671.1	2160655-2162208 m	Similar to putative membrane arginine transporter
SEQ ID n° 6323	4187	2038	SA-670.1	2018869-2019678 m	Similar to mannose-specific PTS enzyme IIC
SEQ ID n° 6324	4188	2039	SA-671.1	2018045-2018872 m	Similar to mannose-specific PTS enzyme IID
SEQ ID n° 6325	4189	2040	SA-672.1	2016258-2017907 m	Similar to two-component sensor histidine kinase
SEQ ID n° 6326	4190	2041	SA-673.1	2015481-2016254 m	Similar to two-component response regulator
SEQ ID n° 6327	4191	2042	SA-674.1	2014432-2015469 m	Similar to iron ABC transporter (binding protein)
SEQ ID n° 6328	4192	2043	SA-675.1	2013713-2014210 p	Similar to unknown proteins
SEQ ID n° 6329	4193	2044	SA-676.1	2012676-2013713 p	Similar to glutamyl-aminopeptidase (hypothetical)
SEQ ID n° 6330	4194	2045	SA-677.1	2012210-2012665 p	Similar to unknown proteins (NrdI)
SEQ ID n° 6331	4195	2046	SA-678.1	2009651-2012053 p	Similar to nucleotidase (esterase), putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 6332	4196	2047	SA-681.1	2162302-2163228 m	Similar to carbamate kinase

SEQ ID n° 6333	4197	2048	SA-680.1	2007175-2009391 m	Similar to GTP pyrophosphokinase (stringent response protein ReA)
SEQ ID n° 6334	4198	2049	SA-681.1	2006722-2007165 m	Similar to unknown proteins
SEQ ID n° 6335	4199	2050	SA-682.1	2005544-2006467 m	Similar to adhesion proteins
SEQ ID n° 6336	4200	2051	SA-686.1	2004072-2005517 m	similar to pneumococcal histidine triad protein B precursor (N-terminal part)
SEQ ID n° 6337	4201	2052	SA-687.1	2002936-2004174 m	similar to pneumococcal histidine triad protein B precursor (C-terminal part)
SEQ ID n° 6338	4202	2053	SA-688.2	2001878-2002654 p	Similar to transcriptional regulator, DeoR family
SEQ ID n° 6339	4203	2054	SA-689.2	712522-713226 p	similar to two-component response regulator
SEQ ID n° 6340	4204	2055	SA-69.1	2163240-2164238 m	similar to ornithine carbamoyltransferase
SEQ ID n° 6341	4205	2056	SA-690.1	710357-712300 p	theronyl-tRNA synthetase
SEQ ID n° 6342	4206	2057	SA-691.1	708566-709900 p	similar to glucosyl transferase
SEQ ID n° 6343	4207	2058	SA-692.1	707566-708564 p	similar to hexosyltransferase
SEQ ID n° 6344	4208	2059	SA-693.1	706055-707521 p	similar to alpha-amylase
SEQ ID n° 6345	4209	2060	SA-694.2	704919-705923 p	catabolite control protein A
SEQ ID n° 6346	4210	2061	SA-696.2	703624-704709 m	similar to X-Pro dipeptidase
SEQ ID n° 6347	4211	2062	SA-697.1	701776-703566 p	similar to beta-N-acetylglucosaminidase
SEQ ID n° 6348	4212	2063	SA-698.1	700948-701760 p	Unknown
SEQ ID n° 6349	4213	2064	SA-699.1	699956-700795 p	similar to oxidoreductase
SEQ ID n° 6350	4214	2066	SA-70.1	2164341-2165636 p	Similar to hypothetical two-component sensor histidine kinase
SEQ ID n° 6351	4215	2067	SA-700.1	698786-699832 p	similar to D-mannonate hydrolase
SEQ ID n° 6352	4216	2068	SA-701.1	697368-698768 p	similar to glucuronate isomerase
SEQ ID n° 6353	4217	2069	SA-702.1	696734-697351 p	similar to 2-dehydro-3-deoxyphosphogluconate aldolase
SEQ ID n° 6354	4218	2070	SA-704.1	695946-696617 p	similar to transcriptional regulator (GntR family)
SEQ ID n° 6355	4219	2071	SA-705.1	694118-695917 p	similar to beta-glucuronidase
SEQ ID n° 6356	4220	2072	SA-706.1	693076-694101 p	similar to 2-keto-3-deoxygluconate kinase
SEQ ID n° 6357	4221	2073	SA-707.1	691459-693009 p	similar to transporter
SEQ ID n° 6358	4222	2074	SA-708.2	690261-691253 p	similar to D-lactate dehydrogenase
SEQ ID n° 6359	4223	2075	SA-71.1	2165633-2166466 p	Similar to hypothetical two-component response regulator
SEQ ID n° 6360	4224	2076	SA-710.2	689189-690247 p	similar to PTS enzyme IIBC
SEQ ID n° 6361	4225	2077	SA-712.1	688112-688996 m	similar to transcriptional regulator, LysR family
SEQ ID n° 6362	4226	2078	SA-713.1	687257-688075 p	Unknown
SEQ ID n° 6363	4227	2079	SA-714.1	686320-687093 p	putative transmembrane protein
SEQ ID n° 6364	4228	2080	SA-715.1	685667-686323 p	similar to ABC transporter (ATP-binding protein)

SEQ ID n° 6365	4229	2081	SA-716.1	684812-685450 m	similar to unknown proteins
SEQ ID n° 6366	4230	2082	SA-717.2	683949-684734 p	similar to competence associated membrane nuclease
SEQ ID n° 6367	4231	2083	SA-719.2	1023804-1025141 m	similar to plasmid proteins
SEQ ID n° 6368	4232	2084	SA-72.1	2166888-2167410 p	Similar to osmoprotectant ABC transporter (ATP-binding protein)
SEQ ID n° 6369	4233	2085	SA-720.1	1023206-1023790 m	Unknown
SEQ ID n° 6370	4234	2086	SA-721.2	1022249-1023067 m	similar to plasmid partition protein ParA
SEQ ID n° 6371	4235	2087	SA-722.2	1021974-1022252 m	Unknown
SEQ ID n° 6372	4236	2088	SA-723.2	1021578-1021961 m	similar to replication initiation protein from Staphylococci plasmid
SEQ ID n° 6373	4237	2089	SA-725.2	1021262-1021573 m	Unknown
SEQ ID n° 6374	4238	2090	SA-726.2	1019800-1021128 m	Unknown
SEQ ID n° 6375	4239	2091	SA-727.1	1018711-1019397 m	similar to unknown protein
SEQ ID n° 6376	4240	2092	SA-728.1	1017948-1018721 m	similar to unknown protein
SEQ ID n° 6377	4241	2093	SA-729.1	1016320-1017915 p	similar to oligopeptide and pheromone binding protein
SEQ ID n° 6378	4242	2096	SA-731.2	1014797-1015867 m	similar to integrase/recombinase
SEQ ID n° 6379	4243	2097	SA-732.1	1013816-1014754 p	similar to two-component sensor histidine kinase
SEQ ID n° 6380	4244	2098	SA-733.1	1013151-1013819 p	similar to two-component response regulator
SEQ ID n° 6381	4245	2099	SA-734.1	1011086-1013041 p	1.2 Transport/binding proteins and lipoproteins
SEQ ID n° 6382	4246	2100	SA-735.1	1010332-1011084 p	similar to ABC transporter (ATP-binding protein)
SEQ ID n° 6383	4247	2101	SA-736.1	1009866-1010306 p	similar to Lactococcus lactis nisin-resistance protein C-terminal part
SEQ ID n° 6384	4248	2102	SA-737.1	1009345-1009884 p	similar to Lactococcus lactis nisin-resistance protein N-terminal part
SEQ ID n° 6385	4249	2103	SA-738.1	1008691-1009179 m	similar to unknown protein
SEQ ID n° 6386	4250	2104	SA-739.1	1007811-1008659 p	similar to other lipoprotein
SEQ ID n° 6387	4251	2105	SA-74.1	2167413-2168927 p	similar to osmoprotectant ABC transporter permease and substrate binding protein
SEQ ID n° 6388	4252	2106	SA-740.1	1006935-1007678 p	similar to unknown protein
SEQ ID n° 6389	4253	2107	SA-741.1	1005523-1006857 p	similar to glucose-inhibited division protein
SEQ ID n° 6390	4254	2108	SA-742.1	1004711-1005409 p	similar to transcriptional regulator (GntR family)
SEQ ID n° 6391	4255	2109	SA-743.2	1002936-1004498 m	similar to GMP synthetase
SEQ ID n° 6392	4256	2110	SA-744.3	439342-440142 p	similar to unknown proteins
SEQ ID n° 6393	4257	2111	SA-745.1	440132-440767 p	similar to unknown protein
SEQ ID n° 6394	4258	2112	SA-746.1	441254-441733 p	similar to unknown proteins

SEQ ID n° 6395	4259	2113	SA-747.1	441769-442920 p	similar to transcription termination-antitermination factor nusA
SEQ ID n° 6396	4260	2114	SA-748.1	442942-443238 p	similar to unknown protein
SEQ ID n° 6397	4261	2115	SA-749.1	443231-443533 p	similar to putative ribosomal protein
SEQ ID n° 6398	4262	2116	SA-75.1	2169006-2171546 m	membrane protein similar to other proteins
SEQ ID n° 6399	4263	2117	SA-751.1	443553-446336 p	initiation factor 2
SEQ ID n° 6400	4264	2118	SA-752.1	446427-446795 p	ribosome binding factor A
SEQ ID n° 6401	4265	2119	SA-753.1	446879-447883 m	similar to esterase
SEQ ID n° 6402	4266	2120	SA-754.1	448047-448463 p	similar to negative transcriptional regulator (copper transport operon)
SEQ ID n° 6403	4267	2121	SA-756.1	448476-450710 p	similar to copper-transporting ATPase (CopA)
SEQ ID n° 6404	4268	2122	SA-757.1	450751-450957 p	similar to copper chaperone (copper transport operon)
SEQ ID n° 6405	4269	2123	SA-758.1	451067-451681 p	similar to unknown protein
SEQ ID n° 6406	4270	2124	SA-759.1	451696-452508 p	similar to unknown protein
SEQ ID n° 6407	4271	2125	SA-76.1	2171530-2172279 m	similar to other proteins
SEQ ID n° 6408	4272	2126	SA-760.1	452621-455263 p	DNA polymerase I
SEQ ID n° 6409	4273	2127	SA-761.1	455293-455733 p	similar to unknown protein
SEQ ID n° 6410	4274	2128	SA-762.1	455815-456294 p	similar to transcription regulator (Fur family)
SEQ ID n° 6411	4275	2129	SA-765.1	456447-458012 p	similar to fibrinogen binding protein, putative peptidoglycan linked protein (LPXTG motif)
SEQ ID n° 6412	4276	2130	SA-766.1	458125-458811 p	similar to two-component response regulator
SEQ ID n° 6413	4277	2131	SA-767.1	458813-459850 p	similar to two-component sensor histidine kinase
SEQ ID n° 6414	4278	2132	SA-768.2	459864-460604 m	similar to unknown protein
SEQ ID n° 6415	4279	2133	SA-769.2	460791-461933 p	similar to tRNA-guanine transglycosylase
SEQ ID n° 6416	4280	2134	SA-77.1	2172407-2172763 m	similar to unknown proteins
SEQ ID n° 6417	4281	2135	SA-770.1	462043-462351 p	similar to unknown protein (putative zinc finger motif)
SEQ ID n° 6418	4282	2136	SA-772.1	462358-462897 p	similar to biotin synthase
SEQ ID n° 6419	4283	2137	SA-773.1	463036-463812 p	similar to unknown protein
SEQ ID n° 6420	4284	2138	SA-774.2	463812-464318 p	similar to unknown protein
SEQ ID n° 6421	4285	2143	SA-78.2	2172831-2175149 m	similar to unknown proteins
SEQ ID n° 6422	4286	2145	SA-782.2	633310-635034 p	similar to negative regulator of FtsZ ring formation protein EzrA
SEQ ID n° 6423	4287	2146	SA-783.1	635128-635769 p	similar to phosphoserine phosphatase
SEQ ID n° 6424	4288	2147	SA-784.1	635790-636275 m	similar to unknown proteins
SEQ ID n° 6425	4289	2148	SA-785.1	636288-636743 m	similar to unknown proteins
SEQ ID n° 6426	4290	2149	SA-786.1	636941-638248 p	enolase

SEQ ID n° 6427	4291	2150	SA-787.1	638356-639420 m	similar to unknown proteins
SEQ ID n° 6428	4292	2151	SA-788.1	639649-640932 p	similar to 5-enolpyruvylshikimate-3-phosphate synthase
SEQ ID n° 6429	4293	2152	SA-789.1	640925-641437 p	similar to shikimate kinase
SEQ ID n° 6430	4294	2154	SA-790.1	641494-642867 p	Similar to membrane bound transcriptional regulator
SEQ ID n° 6431	4295	2155	SA-792.1	642968-644323 p	similar to putative RNA methyltransferase
SEQ ID n° 6432	4296	2156	SA-793.1	644431-644652 p	hypothetical CDS
SEQ ID n° 6433	4297	2157	SA-794.1	644770-645507 p	similar to diadenosine tetraphosphatase, acid phosphatase
SEQ ID n° 6434	4298	2158	SA-795.1	645828-646346 p	similar to unknown proteins
SEQ ID n° 6435	4299	2159	SA-796.1	646475-646711 m	similar to putative transcriptional regulator (TetR/AcrR family) C-terminal part
SEQ ID n° 6436	4300	2160	SA-797.1	646692-647000 m	similar to putative transcriptional regulator (TetR/AcrR family) N-terminal part
SEQ ID n° 6437	4301	2161	SA-798.1	647183-647515 p	similar to C protein alpha-antigen from Streptococcus agalactiae] N-terminal part
SEQ ID n° 6438	4302	2162	SA-799.1	647636-648562 m	similar to transposase, truncated.
SEQ ID n° 6439	4303	2163	SA-8.1	1028864-1031071 m	similar to unknown proteins
SEQ ID n° 6440	4304	2164	SA-80.1	2175288-2175827 p	similar to transcriptional regulator (TetR/AcrR family)
SEQ ID n° 6441	4305	2165	SA-800.1	648505-648780 m	similar to transposase N-terminal part
SEQ ID n° 6442	4306	2166	SA-801.1	648913-649050 p	similar to unknown protein
SEQ ID n° 6443	4307	2167	SA-803.1	649438-649785 m	similar to chaperonin (heat shock protein 33 homolog)
SEQ ID n° 6444	4308	2168	SA-804.1	649979-650398 m	similar to transcriptional regulator (C-terminal part)
SEQ ID n° 6445	4309	2169	SA-805.1	650399-651187 m	similar to transcriptional regulator (N-terminal part)
SEQ ID n° 6446	4310	2170	SA-806.1	651570-653234 p	Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 6447	4311	2171	SA-807.1	653323-654246 p	Putative peptidoglycan bound protein (LPXTG motif)
SEQ ID n° 6448	4312	2172	SA-808.1	654248-655165 p	similar to sortase protein
SEQ ID n° 6449	4313	2173	SA-809.3	655122-655973 p	similar to sortase protein
SEQ ID n° 6450	4314	2174	SA-81.1	2175913-2176209 m	similar to unknown proteins
SEQ ID n° 6451	4315	2176	SA-811.2	2144199-2144348 p	50S Ribosomal protein L33
SEQ ID n° 6452	4316	2177	SA-812.1	2144001-2144183 p	50S Ribosomal protein L32
SEQ ID n° 6453	4317	2178	SA-814.1	2142501-2143781 m	histidyl-tRNA synthetase
SEQ ID n° 6454	4318	2179	SA-817.2	2140657-2142408 m	aspartyl-tRNA synthetase
SEQ ID n° 6455	4319	2180	SA-819.2	2139723-2140667 m	similar to unknown proteins
SEQ ID n° 6456	4320	2181	SA-82.1	2176453-2177064 m	30S ribosomal protein S4
SEQ ID n° 6457	4321	2182	SA-820.1	2138743-2139615 m	similar to unknown proteins
SEQ ID n° 6458	4322	2183	SA-821.1	2138408-2138716 p	similar to unknown proteins
SEQ ID n° 6459	4323	2184	SA-822.1	2136629-2138320 p	arginyl-tRNA synthetase

SEQ ID n° 6460	4324	2185	SA-823.1	2135970-2136416 m	similar to arginine repressor ArgR
SEQ ID n° 6461	4325	2186	SA-825.1	2133337-2135913 m	DNA mismatch repair protein MutS
SEQ ID n° 6462	4326	2187	SA-826.1	2133077-2133280 p	similar to cold shock protein E
SEQ ID n° 6463	4327	2188	SA-827.1	2130878-2132851 m	similar to DNA mismatch repair MutL
SEQ ID n° 6464	4328	2189	SA-828.1	2129614-2130846 m	putative membrane-spanning protein (efflux transporter ?)
SEQ ID n° 6465	4329	2190	SA-829.1	2129022-2129612 m	similar to Holiday junction DNA helicase
SEQ ID n° 6466	4330	2191	SA-83.1	2177394-2177681 m	similar to unknown proteins
SEQ ID n° 6467	4331	2192	SA-831.1	2128448-2128999 m	similar to 3-methyl-adenine DNA glycosylase I
SEQ ID n° 6468	4332	2193	SA-832.1	2127100-2128359 m	similar to competence-damage inducible protein ClnA
SEQ ID n° 6469	4333	2194	SA-833.1	2125887-2127026 m	recombination protein RecA
SEQ ID n° 6470	4334	2195	SA-834.1	2125273-2125671 m	similar to unknown proteins
SEQ ID n° 6471	4335	2196	SA-835.1	2124805-2125071 m	similar to unknown proteins
SEQ ID n° 6472	4336	2197	SA-837.1	2124386-2124805 m	similar to unknown proteins
SEQ ID n° 6473	4337	2198	SA-838.1	2124043-2124360 m	similar to unknown proteins
SEQ ID n° 6474	4338	2199	SA-839.2	2122252-2123793 m	similar to unknown proteins
SEQ ID n° 6475	4339	2200	SA-84.1	2177693-2179048 m	replicative DNA helicase DnaC
SEQ ID n° 6476	4340	2201	SA-842.3	216500-218530 p	Similar to trehalose-specific PTS enzyme IIABC
SEQ ID n° 6477	4341	2202	SA-843.1	218752-220377 p	similar to trehalose-6-phosphate hydrolase
SEQ ID n° 6478	4342	2203	SA-844.1	220597-222633 p	similar to hypothetical transcriptional antiterminator (BglG family)
SEQ ID n° 6479	4343	2204	SA-845.1	222636-222920 p	similar to unknown proteins
SEQ ID n° 6480	4344	2205	SA-846.1	222933-224288 p	putative transmembrane protein similar to unknown proteins
SEQ ID n° 6481	4345	2207	SA-848.1	224291-225148 p	similar to other proteins (including putative transketolase)
SEQ ID n° 6482	4346	2208	SA-849.1	225145-226074 p	similar to other proteins (including putative transketolase)
SEQ ID n° 6483	4347	2209	SA-85.1	2179091-2179543 m	50S ribosomal protein L9
SEQ ID n° 6484	4348	2210	SA-850.1	226102-227442 p	similar to unknown proteins
SEQ ID n° 6485	4349	2211	SA-851.1	227530-227799 p	ribosomal protein S15
SEQ ID n° 6486	4350	2212	SA-852.1	228180-230309 p	polynucleotide phosphorylase, alpha chain
SEQ ID n° 6487	4351	2213	SA-853.1	230311-231063 p	similar to unknown protein
SEQ ID n° 6488	4352	2214	SA-854.1	231072-231656 p	similar to serine acetyltransferase
SEQ ID n° 6489	4353	2215	SA-855.1	231666-231848 p	Unknown
SEQ ID n° 6490	4354	2216	SA-856.1	231845-233188 p	cysteinyI-tRNA synthetase
SEQ ID n° 6491	4355	2217	SA-857.1	233181-233567 p	similar to unknown proteins
SEQ ID n° 6492	4356	2218	SA-858.1	233676-234425 p	similar to tRNA/rRNA methyltransferase
SEQ ID n° 6493	4357	2219	SA-859.1	234422-234940 p	similar to unknown protein
SEQ ID n° 6494	4358	2220	SA-86.1	2179549-2181531 m	similar to unknown proteins

SEQ ID n° 6495	4359	2221	SA-861.1	235033-235893 p	similar to unknown protein
SEQ ID n° 6496	4360	2222	SA-863.3	236878-238059 m	similar to transposase
SEQ ID n° 6497	4361	2223	SA-864.3	181438-182697 m	tyrosyl-tRNA synthetase
SEQ ID n° 6498	4362	2224	SA-866.2	182808-185105 p	similar to penicillin-binding protein 1b
SEQ ID n° 6499	4363	2226	SA-868.1	185629-189204 p	RNA polymerase beta-subunit
SEQ ID n° 6500	4364	2227	SA-87.1	2181598-2183499 m	Similar to GidA protein
SEQ ID n° 6501	4365	2228	SA-872.1	189321-192971 p	RNA polymerase beta -subunit
SEQ ID n° 6502	4366	2229	SA-874.1	193085-193450 p	similar to unknown proteins
SEQ ID n° 6503	4367	2230	SA-875.1	193623-194594 p	similar to transporter (competence protein)
SEQ ID n° 6504	4368	2231	SA-876.1	194440-195531 p	probably part of the DNA transport machinery, ComGB protein
SEQ ID n° 6505	4369	2232	SA-877.1	195528-195857 p	similar to exogenous DNA-binding protein comGC
SEQ ID n° 6506	4370	2233	SA-878.1	195832-196245 p	similar to hypothetical competence proteins
SEQ ID n° 6507	4371	2234	SA-879.1	196217-196516 p	similar to unknown proteins
SEQ ID n° 6508	4372	2235	SA-880.1	196470-196931 p	similar to hypothetical competence proteins
SEQ ID n° 6509	4373	2236	SA-881.1	196909-197280 p	similar to unknown proteins
SEQ ID n° 6510	4374	2237	SA-882.1	197395-198369 p	similar to unknown proteins
SEQ ID n° 6511	4375	2238	SA-883.1	198401-199594 p	acetate kinase
SEQ ID n° 6512	4376	2239	SA-884.1	199745-199951 p	similar to transcriptional regulator
SEQ ID n° 6513	4377	2240	SA-885.2	200188-200643 p	similar to unknown proteins
SEQ ID n° 6514	4378	2241	SA-887.2	579476-579673 p	Unknown
SEQ ID n° 6515	4379	2242	SA-889.1	579717-580649 m	similar to dihydroxotrate dehydrogenase A
SEQ ID n° 6516	4380	2243	SA-89.1	2183669-2184280 m	similar to unknown proteins
SEQ ID n° 6517	4381	2244	SA-890.1	580836-582071 m	similar to Cell Wall Muropeptide Branching Enzyme
SEQ ID n° 6518	4382	2245	SA-891.1	582090-583301 m	similar to cell wall muropeptide branching enzyme
SEQ ID n° 6519	4383	2246	SA-892.1	583314-584534 m	similar to cell wall muropeptide branching enzyme
SEQ ID n° 6520	4384	2247	SA-893.1	584534-585346 m	similar to unknown proteins
SEQ ID n° 6521	4385	2248	SA-894.1	585417-586733 m	similar to unknown proteins
SEQ ID n° 6522	4386	2249	SA-895.1	586809-587195 p	similar to unknown proteins
SEQ ID n° 6523	4387	2250	SA-896.1	587539-590223 p	Similar to cation-transporting P-ATPase
SEQ ID n° 6524	4388	2251	SA-897.1	590268-591128 m	similar to unknown proteins
SEQ ID n° 6525	4389	2252	SA-898.1	591280-593211 p	similar to fructose-1,6-bisphosphatase
SEQ ID n° 6526	4390	2253	SA-899.1	593301-594425 p	Similar to other proteins
SEQ ID n° 6527	4391	2254	SA-9.1	1031191-1031673 m	Unknown
SEQ ID n° 6528	4392	2255	SA-90.1	2184312-2185433 m	similar to tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase

SEQ ID n° 6529	4393	2256	SA-901.1	594579-595592 p	Similar to peptide chain release factor RF-2
SEQ ID n° 6530	4394	2257	SA-902.1	595611-596303 p	Similar to cell division ATP-binding protein FtsE
SEQ ID n° 6531	4395	2258	SA-904.1	596287-597216 p	Similar to cell-division protein FtsX
SEQ ID n° 6532	4396	2259	SA-905.1	597269-597979 m	Similar to unknown proteins
SEQ ID n° 6533	4397	2260	SA-906.2	597976-598611 m	similar to unknown proteins
SEQ ID n° 6534	4398	2261	SA-907.2	909887-910087 p	H+-transporting ATP synthase c chain
SEQ ID n° 6535	4399	2262	SA-908.2	910120-910836 p	H+-transporting ATP synthase a chain
SEQ ID n° 6536	4400	2263	SA-909.1	910854-911351 p	H+-transporting ATP synthase b chain
SEQ ID n° 6537	4401	2264	SA-91.1	2185679-2186347 p	similar to L-serine dehydratase beta subunit
SEQ ID n° 6538	4402	2265	SA-910.1	911351-911887 p	H+-transporting ATP synthase delta chain
SEQ ID n° 6539	4403	2266	SA-911.1	911903-913408 p	H+-transporting ATP synthase alpha chain
SEQ ID n° 6540	4404	2267	SA-912.1	913424-914305 p	H+-transporting ATP synthase gamma chain
SEQ ID n° 6541	4405	2268	SA-913.1	914379-915785 p	H+-transporting ATP synthase beta chain
SEQ ID n° 6542	4406	2269	SA-914.1	915798-916211 p	H+-transporting ATP synthase epsilon chain
SEQ ID n° 6543	4407	2270	SA-916.1	916569-917840 p	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
SEQ ID n° 6544	4408	2271	SA-917.1	918108-918965 p	similar to competence associated membrane nuclease
SEQ ID n° 6545	4409	2272	SA-918.1	919256-920296 p	Phenylalanyl-tRNA synthetase alpha chain
SEQ ID n° 6546	4410	2273	SA-919.1	920379-920900 p	similar to other proteins
SEQ ID n° 6547	4411	2274	SA-920.1	920954-923359 p	Phenylalanyl-tRNA synthetase beta chain
SEQ ID n° 6548	4412	2275	SA-921.1	923428-924096 m	similar to unknown proteins (C-terminal part)
SEQ ID n° 6549	4413	2276	SA-923.1	924207-927440 p	similar to ATP-dependent exonuclease, subunit B
SEQ ID n° 6550	4414	2277	SA-925.1	927430-931053 p	similar to ATP-dependent exonuclease, subunit A
SEQ ID n° 6551	4415	2278	SA-926.1	931066-931992 p	similar to cation transporter
SEQ ID n° 6552	4416	2279	SA-928.2	931967-933343 m	similar to putative tRNA modification GTPase TrmE
SEQ ID n° 6553	4417	2280	SA-93.1	2186362-2187234 p	similar to L-serine dehydratase alpha subunit
SEQ ID n° 6554	4418	2281	SA-930.2	866828-867787 m	Similar to ribonucleotide diphosphate reductase small subunit
SEQ ID n° 6555	4419	2282	SA-931.2	867990-870149 m	Similar to ribonucleotide reductase large subunit
SEQ ID n° 6556	4420	2283	SA-932.2	870227-870451 m	Similar to glutaredoxin
SEQ ID n° 6557	4421	2284	SA-934.1	870833-871096 p	Similar to histidine-containing phosphocarrier protein of the phosphotransferase system (PTS)
SEQ ID n° 6558	4422	2285	SA-936.1	871101-872834 p	phosphoenolpyruvate:sugar phosphotransferase system enzyme
SEQ ID n° 6559	4423	2286	SA-937.1	872984-874411 p	similar to glyceraldehyde-3-phosphate dehydrogenase
SEQ ID n° 6560	4424	2287	SA-939.1	874551-875804 p	similar to oligosaccharide deacetylase
SEQ ID n° 6561	4425	2288	SA-94.1	2187369-2188046 m	similar to other proteins

SEQ ID n° 6562	4426	2289	SA-941.1	875835-876917 m	similar to RNA helicase
SEQ ID n° 6563	4427	2290	SA-942.1	877062-877691 p	similar to uridine kinase
SEQ ID n° 6564	4428	2291	SA-943.1	877778-878275 p	similar to unknown proteins
SEQ ID n° 6565	4429	2292	SA-944.1	878275-879939 p	similar to DNA polymerase III gamma/tau subunit
SEQ ID n° 6566	4430	2293	SA-945.1	880028-880222 p	similar to unknown proteins
SEQ ID n° 6567	4431	2294	SA-946.1	880203-881138 m	similar to transcriptional repressor of the biotin operon
SEQ ID n° 6568	4432	2295	SA-947.1	881323-882519 p	S-adenosylmethionine synthetase
SEQ ID n° 6569	4433	2296	SA-948.1	883038-884945 p	similar to fibronectin-binding protein
SEQ ID n° 6570	4434	2297	SA-949.1	885012-885557 p	Unknown
SEQ ID n° 6571	4435	2298	SA-950.1	885717-885845 p	hypothetical protein
SEQ ID n° 6572	4436	2299	SA-951.1	885953-886519 p	similar to unknown proteins
SEQ ID n° 6573	4437	2300	SA-952.1	886516-887070 p	similar to unknown proteins
SEQ ID n° 6574	4438	2301	SA-953.2	887074-888360 p	similar to cation ABC transporter (ATP-binding protein)
SEQ ID n° 6575	4439	2302	SA-954.1	1170326-1171621 m	Unknown
SEQ ID n° 6576	4440	2303	SA-955.1	1171761-1172060 m	Similar to unknown protein
SEQ ID n° 6577	4441	2304	SA-956.1	1172071-1173318 m	Similar to DNA-methyltransferase
SEQ ID n° 6578	4442	2305	SA-957.1	1173315-1174946 m	Similar to plasmid relaxase and mobilisation protein A
SEQ ID n° 6579	4443	2306	SA-958.1	1174918-1175292 m	similar to unknown proteins
SEQ ID n° 6580	4444	2307	SA-959.1	1175295-1175858 m	Unknown
SEQ ID n° 6581	4445	2308	SA-96.1	2188179-2188718 m	similar to aggregation promoting protein (adhesin involved in high frequency of conjugation)
SEQ ID n° 6582	4446	2309	SA-960.1	1175861-1176157 m	hypothetical CDS
SEQ ID n° 6583	4447	2310	SA-961.1	1176201-1176500 m	Unknown
SEQ ID n° 6584	4448	2311	SA-964.1	1176552-1179788 m	similar to plasmid unknown proteins
SEQ ID n° 6585	4449	2312	SA-966.1	1179790-1180155 m	Unknown
SEQ ID n° 6586	4450	2313	SA-967.1	1180197-1182242 m	Similar to transfer complex protein TrsK - Lactococcus lactis plasmid pMRC01
SEQ ID n° 6587	4451	2314	SA-968.1	1182242-1182733 m	Unknown
SEQ ID n° 6588	4452	2315	SA-969.1	1182755-1183537 m	Similar to gram positive plasmid protein
SEQ ID n° 6589	4453	2316	SA-970.1	1183537-1183809 m	Unknown
SEQ ID n° 6590	4454	2317	SA-971.1	1183829-1184434 m	Unknown
SEQ ID n° 6591	4455	2318	SA-973.1	1184455-1187145 m	similar to plasmid unknown proteins
SEQ ID n° 6592	4456	2319	SA-974.1	1187178-1187708 p	Unknown
SEQ ID n° 6593	4457	2320	SA-975.2	1187902-1190292 m	Similar to plasmid transfer complex protein TrsE
SEQ ID n° 6594	4458	2321	SA-976.2	1899710-1900897 m	similar to two-component sensor histidine kinase
SEQ ID n° 6595	4459	2322	SA-977.1	1898029-1899564 m	Similar to D-alanine-D-alanyl carrier protein ligase

SEQ ID n° 6596	4460	2323	SA-978.1	1896767-1898032 m	similar to LTA D-alanylation protein DltB
SEQ ID n° 6597	4461	2324	SA-98.1	2188934-2189728 m	putative ABC transporter (permease)
SEQ ID n° 6598	4462	2325	SA-980.1	1896513-1896752 m	Similar to D-alanyl carrier protein
SEQ ID n° 6599	4463	2326	SA-981.1	1895258-1896520 m	similar to LTA D-alanine transfer protein DltD
SEQ ID n° 6600	4464	2327	SA-982.1	1894691-1895083 m	Unknown
SEQ ID n° 6601	4465	2328	SA-983.1	1893399-1894691 m	Unknown
SEQ ID n° 6602	4466	2329	SA-984.1	1893016-1893408 m	Unknown
SEQ ID n° 6603	4467	2330	SA-985.1	1892725-1893006 m	Unknown
SEQ ID n° 6604	4468	2331	SA-986.4	1891761-1892543 m	similar to unknown proteins
SEQ ID n° 6605	4469	2332	SA-987.4	1891208-1891774 m	similar to unknown proteins
SEQ ID n° 6606	4470	2333	SA-988.4	1890755-1891204 m	histidine triad family protein
SEQ ID n° 6607	4471	2334	SA-989.1	1889856-1890728 m	Similar to 16 rRNA (adenine-N6,N6-)-dimethyltransferase
SEQ ID n° 6608	4472	2335	SA-99.1	2189721-2190563 m	1.2 Transport/binding proteins and lipoproteins
SEQ ID n° 6609	4473	2336	SA-990.1	1889001-1889852 m	Similar to 23S ribosomal RNA methyltransferase
SEQ ID n° 6610	4474	2337	SA-991.1	1887973-1888845 m	similar to unknown proteins
SEQ ID n° 6611	4475	2338	SA-993.1	1887304-1887966 m	similar to ribulose-5-phosphate 3-epimerase
SEQ ID n° 6612	4476	2339	SA-994.1	1886679-1887311 m	Similar to unknown proteins
SEQ ID n° 6613	4477	2340	SA-995.1	1885403-1886677 m	Similar to unknown proteins
SEQ ID n° 6614	4478	2341	SA-996.1	1884472-1885413 m	similar to unknown proteins
SEQ ID n° 6615	4479	2342	SA-997.2	1883563-1884375 m	similar to transcription repressor of purine operon PurR
SEQ ID n° 6616	4480	2343	SA-998.2	200712-201377 p	similar to unknown proteins
SEQ ID n° 6617	4481	2344	SA-999.1	201398-202168 m	similar to pyrroline-5-carboxylate reductase

TABLEAU 4. Localisation des 139 contigs de séquence SEQ ID No. 1 à SEQ ID No. 139 sur la séquence génomique complète (SEQ ID No. 2345).

SEQ ID	Contig	Position sur génôme complet (= SEQ ID n°2345)		Sens
		position Début	position Fin	(m=minus / p=plus)
SEQ ID n°1	Contig1	1356648	1355991	m
SEQ ID n°2	Contig2	341651	341120	m
SEQ ID n°3	Contig3	476798	476418	m
SEQ ID n°4	Contig4	1475712	1475086	m
SEQ ID n°5	Contig5	784818	784243	m
SEQ ID n°6	Contig6	1756826	1757251	p
SEQ ID n°7	Contig7	1950108	1949648	m
SEQ ID n°8	Contig8	138225	138876	p
SEQ ID n°9	Contig9	2097891	2098283	p
SEQ ID n°10	Contig10	1238491	1237984	m
SEQ ID n°11	Contig11	1882130	1881745	m
SEQ ID n°12	Contig12	1089348	1088935	m
SEQ ID n°13	Contig13	555788	555189	m
SEQ ID n°14	Contig14	2017928	2017437	m
SEQ ID n°15	Contig15	1154094	1154701	p
SEQ ID n°16	Contig16	752647	753091	p
SEQ ID n°17	Contig17	1355561	1355078	m
SEQ ID n°18	Contig18	1255951	1256101	p
SEQ ID n°19	Contig19	792712	793148	p
SEQ ID n°20	Contig20	481787	482228	p
SEQ ID n°22	Contig22	1590263	1590842	p
SEQ ID n°23	Contig23	508269	508918	p
SEQ ID n°24	Contig24	1142198	1142488	p
SEQ ID n°25	Contig25	1982019	1981737	m
SEQ ID n°26	Contig26	119342	119919	p
SEQ ID n°28	Contig28	1124069	1123256	m
SEQ ID n°29	Contig29	266586	266900	p
SEQ ID n°30	Contig30	111013	111623	p
SEQ ID n°31	Contig31	1804173	1804706	p
SEQ ID n°32	Contig32	2170341	2169828	m
SEQ ID n°33	Contig33	1959867	1959394	m
SEQ ID n°34	Contig34	1295529	1294939	m
SEQ ID n°35	Contig35	178592	178071	m
SEQ ID n°36	Contig36	1857103	1856614	m
SEQ ID n°37	Contig37	1063484	1063911	p
SEQ ID n°38	Contig38	198025	197570	m
SEQ ID n°39	Contig39	1486076	1486553	p
SEQ ID n°40	Contig40	2033914	2034352	p
SEQ ID n°41	Contig41	737932	738486	p
SEQ ID n°42	Contig42	729008	728453	m
SEQ ID n°43	Contig43	1671733	1672151	p

SEQ ID n°44	Contig44	1103091	1103644	p
SEQ ID n°45	Contig45	700139	699583	m
SEQ ID n°46	Contig46	207521	206897	m
SEQ ID n°47	Contig47	1064808	1065099	p
SEQ ID n°48	Contig48	1091636	1092281	p
SEQ ID n°49	Contig49	1701764	1700906	m
SEQ ID n°50	Contig50	609072	609590	p
SEQ ID n°51	Contig51	1459271	1458780	m
SEQ ID n°52	Contig52	60603	60154	m
SEQ ID n°53	Contig53	289646	289284	m
SEQ ID n°54	Contig54	1536438	1536058	m
SEQ ID n°55	Contig55	509420	510430	p
SEQ ID n°56	Contig56	1559964	1558709	m
SEQ ID n°58	Contig58	2166712	2165923	m
SEQ ID n°59	Contig59	1919605	1920984	p
SEQ ID n°60	Contig60	962333	960438	m
SEQ ID n°61	Contig61	1363649	1365724	p
SEQ ID n°62	Contig62	1140306	1137284	m
SEQ ID n°63	Contig63	1702242	1706039	p
SEQ ID n°64	Contig64	1490271	1493283	p
SEQ ID n°65	Contig65	783206	785628	p
SEQ ID n°66	Contig66	852318	849615	m
SEQ ID n°67	Contig67	1882303	1880181	m
SEQ ID n°68	Contig68	1614050	1618058	p
SEQ ID n°69	Contig69	1484885	1490042	p
SEQ ID n°70	Contig70	510495	516449	p
SEQ ID n°71	Contig71	125082	121213	m
SEQ ID n°72	Contig72	1557644	1551892	m
SEQ ID n°73	Contig73	145707	143269	m
SEQ ID n°74	Contig74	859105	852465	m
SEQ ID n°75	Contig75	1219383	1215342	m
SEQ ID n°76	Contig76	1091627	1086724	m
SEQ ID n°77	Contig77	1245975	1251984	p
SEQ ID n°78	Contig78	115260	121688	p
SEQ ID n°79	Contig79	1100300	1092624	m
SEQ ID n°80	Contig80	1107948	1100525	m
SEQ ID n°81	Contig81	1245466	1237461	m
SEQ ID n°82	Contig82	2111296	2104033	m
SEQ ID n°83	Contig83	33479	27132	m
SEQ ID n°84	Contig84	1339614	1350526	p
SEQ ID n°85	Contig85	2070423	2058143	m
SEQ ID n°86	Contig86	1462530	1470059	p
SEQ ID n°87	Contig87	526582	517432	m
SEQ ID n°88	Contig88	1484487	1470171	m
SEQ ID n°89	Contig89	1124087	1136746	p
SEQ ID n°90	Contig90	1879890	1866931	m
SEQ ID n°91	Contig91	1721684	1706045	m
SEQ ID n°92	Contig92	1358184	1357897	m
SEQ ID n°93	Contig93	1577596	1560798	m
SEQ ID n°94	Contig94	115130	103188	m

SEQ ID n°95	Contig95	1921051	1933881	p
SEQ ID n°96	Contig96	1944905	1933782	m
SEQ ID n°97	Contig97	1919624	1906953	m
SEQ ID n°98	Contig98	2090559	2103658	p
SEQ ID n°99	Contig99	1237482	1219423	m
SEQ ID n°100	Contig100	1123110	1108191	m
SEQ ID n°101	Contig101	1551836	1529458	m
SEQ ID n°102	Contig102	1818811	1800978	m
SEQ ID n°103	Contig103	764781	783195	p
SEQ ID n°104	Contig104	1086606	1065938	m
SEQ ID n°105	Contig105	125425	143102	p
SEQ ID n°106	Contig106	962438	984387	p
SEQ ID n°107	Contig107	1169838	1190193	p
SEQ ID n°108	Contig108	2090426	2070667	m
SEQ ID n°109	Contig109	1140315	1169462	p
SEQ ID n°111	Contig111	238297	258413	p
SEQ ID n°112	Contig112	216686	237881	p
SEQ ID n°113	Contig113	2209521	16967	m
SEQ ID n°114	Contig114	1883537	1906918	p
SEQ ID n°115	Contig115	145772	172009	p
SEQ ID n°116	Contig116	508181	477405	m
SEQ ID n°117	Contig117	859233	888273	p
SEQ ID n°118	Contig118	1529046	1494213	m
SEQ ID n°119	Contig119	473132	438871	m
SEQ ID n°120	Contig120	1981657	1945366	m
SEQ ID n°121	Contig121	1613824	1577594	m
SEQ ID n°122	Contig122	1765846	1800817	p
SEQ ID n°123	Contig123	2111499	2153851	p
SEQ ID n°124	Contig124	1721668	1765765	p
SEQ ID n°125	Contig125	984406	1025178	p
SEQ ID n°126	Contig126	1293488	1339586	p
SEQ ID n°127	Contig127	216691	176332	m
SEQ ID n°128	Contig128	1818941	1866861	p
SEQ ID n°129	Contig129	849565	785796	m
SEQ ID n°130	Contig130	888292	960270	p
SEQ ID n°131	Contig131	2208563	2155215	m
SEQ ID n°132	Contig132	33590	88257	p
SEQ ID n°133	Contig133	1982609	2057812	p
SEQ ID n°134	Contig134	1700642	1618142	m
SEQ ID n°135	Contig135	1293063	1190375	m
SEQ ID n°136	Contig136	1366980	1462324	p
SEQ ID n°137	Contig137	390853	434186	p
SEQ ID n°138	Contig138	357393	259739	m
SEQ ID n°139	Contig139	527049	716899	p

TABLEAU 5. Propriété d'adhérence à des cellules épithéliales humaines en culture de la souche NEM316 de *S. agalactiae* et de souches mutantes dérivées.

Souche	Gène inactivé	% d'adhésion ^a	
		Cellules A549	Cellules Hela
NEM316	aucun	9	16
NEM1979	IPF N° 1268 (<i>srtA</i>)	1,5	2
NEM2056	IPF N° 678	2	n.t.
NEM2057	IPF N° 1503	4,5	n.t.

- 5 ^a, le pourcentage d'adhésion correspond au nombre de bactéries (Unité Formant des Colonies, UFC) restant adhérentes aux cellules après lavage avec du tampon PBS par rapport au nombre d'UFC ajoutées à la monocouche de cellules épithéliales.

TABLEAU 6. Gènes de la souche de *S. agalactiae* NEM316 codant pour des protéines de surface avec un motif d'ancrage LPXTG^a

Seq ID (ADN)	IPF N°	taille pbases	Site de coupure	Protéines homologues	% d'acides aminés identiques (similaires) / longueur de la région similaire ^b	Fonction prédite ^c
6194	523	1126	LPXT/G	Alp2 (<i>S. agalactiae</i>) Alp3 (<i>S. agalactiae</i>) R28 (<i>S. pyogenes</i>)	74 (77) / 798 71 (76) / 877 69 (75) / 1103	inconnue
6236	571	1310	"	Hsa (<i>S. gordonii</i>) SrpA (<i>S. cristatus</i>)	50 (60) / 1314 43 (53) / 1248	Protéine liant l'acide sialique
5497	220	1634	"	Ssp-5 (<i>S. gordonii</i>), PAa (<i>S. intermedius</i>)	30 (43) / 1385 31 (45) / 1285	Protéine liant l'acide sialique
5491	2192	512	"	EaeH (<i>E. coli</i> O157:H7)	25 (38) / 358	Adhésion
5103	1716	643	"	M-like protein (<i>S. equi</i>) PspC (<i>S. pneumoniae</i>)	31 (46) / 302 23 (38) / 795	inconnue Adhésion
4705	1247	932	"	SpaA (<i>S. sobrinus</i>)	38(52) / 406	inconnue
5610	2337	308	"	No homology in public databases		inconnue
5234	1861	543	"	Cell surface protein (<i>S. mutans</i>) CbpD (<i>S. pneumoniae</i>)	50(62) / 183 30(60) / 220	inconnue Protéine liant la choline
4926	1503	1570	"	PrtS (<i>S. thermophilus</i>)	49 (65) / 1596	Serine proteinase
6331	678	800	"	CpdB (<i>S. dysgalactiae</i>) YfKN (<i>Bacillus subtilis</i>)	57(70) / 694 47(66) / 630	Cyclo-nucleotide phosphodiesterase

6247	585	680		AmiC (<i>S. pyogenes</i>)	36 (54) / 478	Amidase
5842	280	1252	"	YbgE (<i>L. lactis</i>)	35 (54) / 492	Amylopullulanase alkaline
5741	2495	410	"	PulA (<i>S. pyogenes</i>)	65 (79) / 1095	inconnue
4921	15	933	"	CG15040 gene product <i>Drosophila melanogaster</i>	23(47) / 373	inconnue
5090	17	240	"	Antigen p200 (<i>Babesia bigemina</i>)	26(50) / 273	inconnue
5180	18	753	"	SpaA (<i>S. sobrinus</i>)	37(52) / 405	inconnue
4706	1248	236	"	Pas (<i>S. intermedius</i>)	36(52) / 399	inconnue
4708	1250	743	"	Plasmid-encoded protein (<i>E. faecalis</i>)	33 (49) / 225	inconnue
5677	2414	253	"	Sec10 (<i>E. faecalis</i>)	24 (37) / 715	Surface exclusion protein
6246	584*	192	"	Plasmid-encoded protein (<i>E. faecalis</i>)	31 (47) / 263	inconnue
6411	765	521	"	No homology in public databases	22 (40) / 784	Surface exclusion protein
5578	2300	901	IPXT/G	No homology in public databases	33 (47) / 211	inconnue
6446	806	554	"	Hypothetical protein 2 (<i>Lactobacillus leichmannii</i>)	32(46) / 176	Protéine liant la fibronectine
6447	807	307	"	Fimbrial structural subunit (<i>Actinomyces naeslundii</i>)	27(42) / 512	inconnue
5607	2334	674	"	No homology in public databases	25(38) / 577	inconnue
				No homology in public databases		inconnue

6209	538	1055	LPXT/S	SPy0843 (<i>S. pyogenes</i>) BspA (<i>Bacteroides forsythus</i>)	72 (81) / 1050 24 (41) / 566	inconnue inconnue
6215	547	1233	"	ScpB (<i>S. agalactiae</i>)	38(55) / 1194	Protéase à sérine
5406	2082	1150	LPXT/N	ScpB (<i>S. agalactiae</i>)	99(99) / 1150	C5a peptidase
5658	2390	690	"	SPy0872 (<i>S. pyogenes</i>)	60(74) / 688	5'-nucleotidase secretée
4965	1551	890	FPKT/G	No homology in public databases		inconnue

^a , Les protéines ancrées au peptidoglycane ont été identifiées par la recherche d'un motif LPXTG ou un motif voisin C-terminal suivi d'un domaine hydrophobe et d'acides aminés basiques. Les similarités par BLASTP avec des protéines à domaine LPXTG connues ont aussi été utilisées.

^b , Seulement les similarité avec une probabilité BLASTP <10⁻¹⁰ ont été considérées comme significatives. ^c , La fonction a été prédite par analogie avec celles des protéines homologues contenues dans la bases de séquences protéique nrprot du NCBI.

TABLEAU 7. Distribution des gènes codant pour des protéines de surface à motif LPXTG parmi des isolats cliniques indépendants de 5 sérotypes de *S. agalactiae*.

Seq ID (ADN)	N° d'IPF	Proportion des souches portant le gène						
		Type Ia 23 isolats	Type Ib 7 isolats	Type II 12 isolats	Type III 39 isolats	Type V 16 isolats	non groupé 2 isolats	Fréquence totale 99 isolats
4926	1503	100	100	100	100	100	100	100
6331	678	100	100	100	100	100	100	100
5491	2192	100	100	100	100	100	100	100
5234	1861	100	100	100	100	100	100	100
6246	584	100	100	100	100	100	100	100
5842	280	100	100	100	100	100	100	100
6247	585	86,9	85,7	91,7	92,3	81,3	100	88,9
4965	1551	34,8	85,7	100	94,8	93,8	100	80,8
6447	807	26,1	85,7	83,3	92,3	93,8	100	79,8
5610	2337	86,9	100	100	56,4	100	50	78,8
6236	571	82,6	100	83,3	46,1	100	50	71,7
5103	1716	91,0	100	83,3	38,5	81,3	50	67,7
5607	2334	56,5	100	83,3	46,1	100	50	65,6
6411	765	30,4	71,4	41,7	61,5	87,5	100	57,6
4921	15	65,2	0	41,7	59	12,5	50	46,5
5090	17	4,3	14,3	16,7	20,5	6,3	0	13,1
5180	18	4,3	14,3	0	20,5	6,3	0	11,1
4706	1248	0	0	0	12,8	6,3	0	6
5497	220	0	14,3	8,3	5,1	0	0	4,5
4708	1250	0	14,3	0	5,1	0	0	3
5677	2414	0	0	0	7,7	0	0	3

TABLEAU 8. Lipoprotéines

Seq ID (ADN)	N° d'IPF	Annotation
6527	9	unknown
6030	339	similar to unknown proteins
6035	344	Similar to ABC transporter (binding protein)
6137	460	Similar to ABC transporter (binding protein)
6178	504	similar to unknown proteins
6294	638	Similar to unknown proteins
6335	682	Similar to adhesion proteins
6377	729	similar to oligopeptide and pheromone binding protein
6386	739	similar to other lipoprotein
4495	1018	Similar to (oligopeptide) ABC transporter (binding protein)
4596	1119	similar to ribose ABC transporter (binding protein)
4636	1162	similar to (amino acid ?) ABC transporter (binding protein)
4730	1280	similar to ABC transporter (binding protein)
4816	1377	Similar to nickel ABC transporter (binding protein)
4836	1399	similar to phosphate ABC transporter (binding protein)
4906	1481	Similar to D,D-carboxypeptidase
4920	1499	similar to peptidyl-prolyl cis-trans isomerase
4925	1502	similar to metal ABC transporter (binding protein)
4963	1547	Unknown
5021	1617	Similar to unknown lipoprotein
5158	1775	similar to ferrichrome ABC transporter (binding protein)
5247	1879	similar to oligopeptide ABC transporter (binding protein)
5306	1955	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)
5417	2099	similar to putative ABC transporter (binding protein)
5423	2103	Unknwon, similar to unknown protein and to B. subtilis SpoIIJ protein
5450	2133	laminin-binding surface protein
5486	2185	putative ABC transporter (binding protein)

5559	2278	putative ABC transporter (binding protein)
5591	2314	similar to protease maturation protein
5677	2414	Similar to plasmid related proteins, Putative peptidoglycan bound protein (LPXTG motif)
5718	2464	Similar to unknown proteins
5732	2482	similar to ferrichrome ABC transporter (binding protein)
5799	2597	similar to amino acid ABC transporter (binding protein)
5800	2598	similar to phosphate ABC transporter (binding protein)
5837	2789	Unknown
5861	2843	Similar to amino acid ABC transporter (binding protein)
5883	2875	Unknown
5923	2922	Similar to amino acid ABC transporter (binding protein)

- Les gènes codant pour des lipoprotéines ont été identifiés sur la base de la prédiction du motif de coupure / modification de type lipoprotéine [S. Hayashi, H. C. Wu. *J Bioenerg Biomembr.* 22, 451 (1990)] et d'un peptide signal (identifié en utilisant
- 5 SignalP vs2.0 [H Nielsen, *Prot Engin* 12, 13-9. (1999)]) et par l'analyse des résultats de comparaison sur les banques de séquence protéiques en utilisant BLAST [S. F. Altschul et al., *Nucleic Acids Res* 25, 3389-402. (1997)].

Tableau 9. Autres protéines de surface

10

Seq ID (ADN)	N° d'IPF	Annotation
6037	347	group B streptococcal surface immunogenic protein
4972	1562	Putative cell wall protein, weakly similar to peptidase or esterase
6569	948	similar to fibronectin-binding protein
5234	1861	similar to cell wall proteins
5530	2238	possible surface protein
5223	1847	CAMP factor

Ces protéines ont été identifiées sur la base de similarité avec d'autres protéines de surface bactérienne et la prédiction d'un peptide signal et ne faisant pas partie des classes des protéines ancrées au peptidoglycane et lipoprotéines.

5 TABLEAU 10. Protéines impliquées dans la biosynthèse de composés polysaccharidiques de la paroi de *S. agalactiae*.

Seq ID (ADN)	N° d'IPF	Annotation	Commentaires
4861	1430	similar to Streptococcus mutans RgpG protein required for biosynthesis of rhamnose-glucose polysaccharide	
6214	544	similar to rhamnosyltransferase	
6061	381	Unnown, Similar to UDP-N-acetylmuramoylalanine--D-glutamate ligase	
6517	890	similar to Cell Wall Muropeptide Branching Enzyme	
6518	891	similar to cell wall muropeptide branching enzyme	
6519	892	similar to cell wall muropeptide branching enzyme	
4743	1295	similar to glycosyltransferases	
6343	692	similar to hexosyltransferase	
6342	691	similar to glucosyl transferase	
5326	1977	Similar to UDP-D-glucose:galactosyl glucosyltransferase	
4952	1532	similar to N-acetylneuraminic acid synthetase	
5619	2346	capsular polysaccharide biosynthesis protein	Biosynthèse de la capsule
5618	2345	similar to glycosyl transferase	
5617	2344	similar to glycosyl transferase	
5616	2343	capsular polysaccharide repeating-unit polymerase	
5615	2342	beta-1,4-galactosyltransferase	
5614	2341	beta-1,4-galactosyltransferase enhancer	
5613	2340	similar to glucose-1-phosphate transferase	
5611	2339	capsular polysaccharide chain length regulator/exporter	
5696	2437	putative chain length regulator CpsC	
5971	301	similar to dTDP-glucose-4,6-dehydratase	
5233	1860	similar to to Cell Wall Muropeptide Branching Enzyme	

5602	2329	similar to capsular polyglutamate biosynthesis	
5156	1773	Similar to UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate ligase	
5574	2297	Similar to putative hexosyltransferase	
5573	2296	Similar to rhamnosyl transferase I	
5654	2386	Similar to capsular polysaccharide synthesis protein	
5656	2388	Similar to putative rhamnosyltransferase	
5526	2233	Similar to putative rhamnosyltransferase	
5527	2235	Similar to nucleotide-sugar dehydratase	
5529	2237	Similar to Pneumococcal LicD2 protein involved in phosphorylcholine metabolism	
5534	2241	similar to rhamnosyltransferase	
5625	2354	similar to putative rhamnosyltransferase	
5626	2355	dTDP-L-rhamnose synthase	
6223	555	Similar to putative glucosyl transferase	
6229	562	Similar to hypothetical glycosyl transferase	
6230	563	Similar to putative glycosyltransferase	
6231	565	Similar to putative glycosyl transferase	
6232	566	Similar to putative glycosyl transferase	
6233	567	Similar to putative glycosyl transferase	
5764	2518	similar to putative sugar transferase	
6095	416	similar to UDP-N-acetylglucosamine pyrophosphorylase	
5089	1699	Similar to UDP-N-acetylmuramate-alanine ligase	
5466	2158	similar to glycosyl transferase	
5465	2157	similar to glycosyl transferase	

Ces gènes ont été identifiés par l'analyse des résultats de similarité avec les séquences protéiques connues en utilisant le logiciel BLASTP. Les produits de ces gènes pourraient intervenir dans la biosynthèse de polysaccharides qui pourraient être

5 des constituants de préparations vaccinales.

REFERENCES

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, [Review] [90 refs] *Nucleic Acids Research*, 25:3389-402.
- 5 2. Birnboim, H. C., 1983, A rapid alkaline extraction method for the isolation of plasmid DNA, *Methods Enzymol.*, 100:243-255.
3. Brodeur, B.B., M. Boyer, I. Charlebois, J. Hamel, F. Couture, C.R. Rioux, and D. Martin, 2000, Identification of Group B Strptococcal Sip Protein, which elicits cross-protective immunity, *Infect. Immun.*, 68:5610-5618.
- 10 4. Buchrieser, C., C. Rusniok, L. Frangeul, E. Couvé, A. Billault, F. Kunst, E. Carniel, and P. Glaser, 1999, The 102 kb locus of *Yersinia pestis* : sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains, *Infect. Immun.*, 67:4851-4861.
5. Ewing, B., and P. Green, 1998, Base-calling of automated sequencer traces using phred. II. Error probabilities, *Genome Res.*, 8:186-194.
- 15 6. Fitch, W.S., 1970, Distinguishing homologous from analogous proteins, *Syst. Zool.*, 19:99-113.
7. Frangeul, L., K.E. Nelson, C. Bushrieser, A. Danchin, P. Glaser, and K.F., 1999, Cloning and assembly strategies in microbial genome projects, *Microbiology*, 20 145:2625-2634.
8. Gordon, D., C. Abajian, and P. Green, 1998, Consed: a graphical tool for sequence finishing, *Genome Res.*, 8:195-202.
10. Li, P., K.C. Kupfer, C.J. Davies, D. Burbee, G.A. Evans, and H.R. Garner, 1997, PRIMO: A primer design program that applies base quality statistics for automated large-scale DNA sequencing, *Genomics*, 40:476-485.
- 25 11. Lukashin, A.V., and M. Borodovsky, 1998, GeneMark.hmm: new solutions for gene finding, *Nucleic Acids Res.* 15:1107-1115.

REVENDICATIONS

1. Séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi les séquences SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345.
2. Séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi :
- a) une séquence nucléotidique comportant au moins 75 % d'identité avec une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345 ;
 - 10 b) une séquence nucléotidique hybridant dans des conditions de forte stringence avec une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, et comprenant au moins 20 nucléotides ;
 - c) une séquence nucléotidique complémentaire d'une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, ou complémentaire d'une séquence
15 nucléotidique telle que définie en a), ou b), ou une séquence nucléotidique de l'ARN correspondant à l'une des séquences a) ou b) ;
 - d) une séquence nucléotidique d'un fragment représentatif d'une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, ou d'un fragment représentatif d'une séquence nucléotidique telle que définie en a), b) ou c) et
20 comprenant au moins 20 nucléotides ;
 - e) une séquence nucléotidique comprenant une séquence telle que définie en a), b), c) ou d) ; et
 - f) une séquence nucléotidique telle que définie en a), b), c), d) ou e) modifiée et comportant au plus 10 % de nucléotides modifiés par rapport à la séquence de
25 référence.
3. Séquence nucléotidique selon la revendication 2, caractérisée en ce qu'il s'agit d'une séquence issue d'une séquence choisie parmi SEQ ID No. 1 à SEQ ID No. 139 et SEQ ID No. 2345, et en ce qu'elle code pour un polypeptide, choisi de préférence parmi les séquences SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481.
4. Séquence nucléotidique caractérisée en ce qu'elle comprend une séquence nucléotidique choisie parmi :
- a) une séquence nucléotidique selon la revendication 3 ou choisie parmi les séquences SEQ ID No. 4482 à SEQ ID No. 6617 ;

- b) une séquence nucléotidique comportant au moins 75 % d'identité avec une séquence nucléotidique selon la revendication 3 ;
- c) une séquence nucléotidique s'hybridant dans des conditions de forte stringence avec une séquence nucléotidique selon la revendication 3 et comprenant au moins 20
5 nucléotides ;
- d) une séquence nucléotidique complémentaire ou d'ARN correspondant à une séquence telle que définie en a), b) ou c) ;
- e) une séquence nucléotidique d'un fragment représentatif d'une séquence telle que définie en a), b), c) ou d) et comprenant au moins 20 nucléotides ; et
- 10 f) une séquence telle que définie en a), b), c), d) ou e) modifiée et comportant au plus 10 % de nucléotides modifiés par rapport à la séquence de référence.

5. Polypeptide codé par une séquence nucléotidique selon l'une des revendications 2 à 4.

6. Polypeptide selon la revendication 5, caractérisé en ce qu'il est choisi parmi
15 les polypeptides choisis parmi SEQ ID No. 140 à SEQ ID No. 2344, et SEQ ID No. 2346 à SEQ ID No. 4481.

7. Polypeptide caractérisé en ce qu'il comprend un polypeptide choisi parmi :

- a) un polypeptide selon l'une des revendications 5 et 6 ;
- b) un polypeptide présentant au moins 80 % d'identité avec un polypeptide selon l'une
20 des revendications 5 et 6 ;
- c) un fragment d'au moins 5 acides aminés d'un polypeptide selon l'une des revendications 5 et 6, ou tel que défini en b) ;
- d) un fragment biologiquement actif d'un polypeptide selon l'une des revendications 5 et 6, ou tel que défini en b) ou c) ; et
- 25 e) un polypeptide selon l'une des revendications 5 et 6 ou tel que défini en b), c) ou d) modifié et comportant au plus 10 % d'acides aminés modifiés par rapport à la séquence de référence.

8. Séquence nucléotidique codant pour un polypeptide selon la revendication
7.

- 30 9. Séquence nucléotidique isolée codant pour un polypeptide spécifique de *Streptococcus agalactiae* choisi parmi les polypeptides de séquence SEQ ID No. 140 à SEQ ID No. 2344 et SEQ ID No. 2346 à SEQ ID No. 4481.

10. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans la biosynthèse des acides aminés ou l'un de ses fragments.

5 11. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans la biosynthèse des cofacteurs, groupes prosthétiques et transporteurs ou l'un de ses fragments.

10 12. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide d'enveloppe cellulaire ou situé à la surface de *Streptococcus agalactiae* ou l'un de ses fragments.

13. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans la machinerie cellulaire ou l'un de ses fragments.

15 14. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme intermédiaire central ou l'un de ses fragments.

15. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme énergétique ou l'un de ses fragments.

20 16. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme des acides gras et des phospholipides ou l'un de ses fragments.

25 17. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme des nucléotides, des purines, des pyrimidines ou nucléosides ou l'un de ses fragments.

30 18. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans les fonctions de régulation ou l'un de ses fragments.

19. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de réplication ou l'un de ses fragments.

20. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de transcription ou l'un de ses fragments.

5 21. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de traduction ou l'un de ses fragments.

22. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de transport et de liaison des protéines ou l'un de ses
10 fragments.

23. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans l'adaptation aux conditions atypiques ou l'un de ses fragments.

24. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9,
15 caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans la sensibilité aux médicaments et analogues ou l'un de ses fragments.

25. Séquence nucléotidique selon l'une des revendications 2 à 4, 8 et 9, caractérisée en ce qu'elle code pour un polypeptide de *Streptococcus agalactiae* impliqué dans les fonctions relatives aux transposons ou l'un de ses fragments.

20 26. Séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi :

a) une séquence choisie parmi les séquences SEQ ID N° 6194,6236,5497,5791,5103,4705,5610,5234,4926,6331,6247,5842,5741,4921,5090,5180,4706,4708,5677,6246,6411,5578,6446,6447,5607,6209,6215,5406,5658,4965, de
25 préférence parmi les séquences SEQ ID N°4926,6331,5491,5234,6246,5842 ;

b) une séquence nucléotidique comportant au moins 75 % d'identité avec une séquence nucléotidique du a) ;

c) une séquence nucléotidique s'hybridant dans des conditions de forte stringence avec une séquence nucléotidique du a) ou b) et comprenant au moins 20
30 nucléotides ;

d) une séquence nucléotidique complémentaire ou d'ARN correspondant à une séquence telle que définie en a), b) ou c) ;

e) une séquence nucléotidique d'un fragment représentatif d'une séquence telle que définie en a), b), c) ou d) et comprenant au moins 20 nucléotides ; et

f) une séquence telle que définie en a), b), c), d) ou e) modifiée et comportant au plus 10 % de nucléotides modifiés par rapport à la séquence de référence ;

et en ce qu'elle code pour une protéine de surface avec un motif d'ancrage LPXTG.

5 27. Séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi les séquences SEQ ID N°6035,6137,6335,6377,6386,4495,4596,4636,4730,4816,4836,4906,4920,4925,5158,5247, 5306,5417,5450,5486,5559,5591,5677,5732,5799,5800,5861,5923 et en ce qu'elle code pour une lipoprotéine.

10 28. Séquence nucléotidique isolée de *Streptococcus agalactiae*, caractérisée en ce qu'elle est choisie parmi les séquences SEQ ID N°4861,6214,6061,6517,6518,6519,4743,6343,6342,5326,4952,5619,5618,5617,5616,5615,5614,5613,5611,5696,5971,5233,5602,5156,5574,5573,5654,5656,5526,5527,5529,5534,5625,5626,6223,6229,6230,6231,6232,6233,5764,6095,5089,5466,5465 et en ce
15 qu'elle code pour une protéine impliquée dans la biosynthèse de composés polysaccharidiques de paroi.

29. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans la biosynthèse des acides aminés ou l'un de ses fragments.

20 30. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans la biosynthèse des cofacteurs, groupes prosthétiques et transporteurs ou l'un de ses fragments.

31. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide d'enveloppe cellulaire ou situé à la surface de *Streptococcus*
25 *agalactiae* ou l'un de ses fragments.

32. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans la machinerie cellulaire ou l'un de ses fragments.

33. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il
30 s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme intermédiaire central ou l'un de ses fragments.

34. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme énergétique ou l'un de ses fragments.

35. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme des acides gras et des phospholipides ou l'un de ses fragments.

5 36. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le métabolisme des nucléotides, des purines, des pyrimidines ou nucléosides ou l'un de ses fragments.

37. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans les fonctions de régulation ou l'un de ses fragments.

10 38. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de réplication ou l'un de ses fragments.

39. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de
15 transcription ou l'un de ses fragments.

40. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de traduction ou l'un de ses fragments.

41. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il
20 s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans le processus de transport et de liaison des protéines ou l'un de ses fragments.

42. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans l'adaptation aux conditions atypiques ou l'un de ses fragments.

25 43. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans la sensibilité aux médicaments et analogues ou l'un de ses fragments.

44. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il s'agit d'un polypeptide de *Streptococcus agalactiae* impliqué dans les fonctions
30 relatives aux transposons ou l'un de ses fragments.

45. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il est codé par une séquence selon la revendication 26 et est une protéine de surface avec un motif d'ancrage LPXTG.

46. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il est codé par une séquence selon la revendication 27 et est une lipoprotéine.
47. Polypeptide selon l'une des revendications 5 à 7, caractérisé en ce qu'il est codé par une séquence selon la revendication 28 et est une protéine impliquée dans la biosynthèse de composés polysaccharidiques de paroi.
48. Séquence nucléotidique utilisable comme amorce ou comme sonde, caractérisée en ce que ladite séquence est choisie parmi les séquences nucléotidiques selon l'une des revendications 2 à 4, 8 à 28.
49. Séquence nucléotidique selon la revendication 48, caractérisée en ce qu'elle est marquée par un composé radioactif ou par un composé non radioactif.
50. Séquence nucléotidique selon l'une des revendications 48 et 49, caractérisée en ce qu'elle est immobilisée sur un support, de manière covalente ou non-covalente.
51. Séquence nucléotidique selon la revendication 50, caractérisée en ce qu'elle est immobilisée sur un support tel qu'un filtre à haute densité ou une puce à ADN.
52. Séquence nucléotidique selon l'une des revendications 49 à 51 pour la détection et/ou l'amplification de séquences nucléiques.
53. Puce à ADN ou filtre, caractérisée en ce qu'elle contient au moins une séquence nucléotidique selon la revendication 51.
54. Puce à ADN ou filtre selon la revendication 53, caractérisée en ce qu'elle contient en outre au moins une séquence nucléotidique d'un micro-organisme autre que *Streptococcus agalactiae*, immobilisée sur le support de ladite puce.
55. Puce à ADN ou filtre selon la revendication 54, caractérisée en ce que le micro-organisme autre est choisi parmi un micro-organisme associé à *Streptococcus agalactiae*, une bactérie du genre *Streptococcus*, et un variant de *Streptococcus agalactiae*.
56. Kit ou nécessaire pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé, caractérisé en ce qu'il comprend une puce à ADN ou un filtre selon la revendication 53.
57. Kit ou nécessaire pour la détection et/ou l'identification d'un micro-organisme, caractérisé en ce qu'il comprend une puce à ADN ou un filtre selon l'une des revendications 54 et 55.
58. Kit ou nécessaire pour la détection et/ou la quantification de l'expression d'au moins un gène de *Streptococcus agalactiae*, caractérisé en ce qu'il comprend une puce à ADN ou un filtre selon l'une des revendications 53 à 55.

59. Vecteur de clonage, et/ou d'expression, caractérisé en ce qu'il contient une séquence nucléotidique selon l'une des revendications 1 à 4, 8 à 28.

60. Cellule hôte, caractérisée en ce qu'elle est transformée par un vecteur selon la revendication 59.

5 61. Cellule hôte selon la revendication 60, caractérisée en ce qu'il s'agit d'une bactérie appartenant au genre *Streptococcus*.

62. Cellule hôte selon la revendication 61, caractérisée en ce qu'il s'agit d'une bactérie appartenant à l'espèce *Streptococcus agalactiae*.

10 63. Végétal ou animal, excepté l'Homme, comprenant une cellule transformée selon l'une des revendications 60 à 62.

64. Procédé de préparation d'un polypeptide, caractérisé en ce que l'on cultive une cellule transformée par un vecteur selon la revendication 59 dans des conditions permettant l'expression dudit polypeptide et que l'on recupère ledit polypeptide recombinant.

15 65. Polypeptide recombinant susceptible d'être obtenu par un procédé selon la revendication 64.

66. Procédé de préparation d'un polypeptide synthétique selon l'une des revendications 5 à 7, 29 à 47, caractérisé en ce que l'on effectue une synthèse chimique dudit polypeptide.

20 67. Polypeptide hybride, caractérisé en ce qu'il comprend au moins la séquence d'un polypeptide selon l'une des revendications 5 à 7, 29 à 47 et 65, et une séquence d'un polypeptide susceptible d'induire une réponse immunitaire chez l'homme ou l'animal.

25 68. Séquence nucléotidique codant pour un polypeptide hybride selon la revendication 67.

69. Vecteur caractérisé en ce qu'il contient une séquence nucléotidique selon la revendication 68.

30 70. Anticorps monoclonal ou polyclonal, ses fragments, ou anticorps chimérique, caractérisé en ce qu'il est capable de reconnaître spécifiquement un polypeptide selon l'une des revendications 5 à 7, 29 à 47, 65 et 67.

71. Anticorps selon la revendication 70, caractérisé en ce qu'il s'agit d'un anticorps marqué.

72 Procédé pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- 5 a) mise en contact de l'échantillon biologique avec un anticorps selon l'une des revendications 70 et 71 ;
b) mise en évidence du complexe antigène-anticorps éventuellement formé.

73. Procédé pour la détection de l'expression d'un gène de *Streptococcus agalactiae* caractérisé en ce que l'on met en contact une souche de *Streptococcus agalactiae*, avec un anticorps selon la revendication 70 ou 71 et que l'on détecte le
10 complexe antigène/anticorps éventuellement formé.

74. Kit ou nécessaire pour la mise en œuvre d'un procédé selon la revendication 72 ou 73, caractérisé en ce qu'il comprend les éléments suivants :

- 15 a) un anticorps selon l'une des revendications 70 et 71 ;
b) éventuellement, les réactifs pour la constitution du milieu propice à la réaction immunologique ;
c) éventuellement, les réactifs permettant la mise en évidence des complexes antigène-anticorps produits par la réaction immunologique.

75. Polypeptide selon l'une des revendications 5 à 7, 32 à 47, 65 et 67, ou anticorps selon l'une des revendications 64 et 65, caractérisé en ce qu'il est immobilisé
20 sur un support, notamment une puce à protéine.

76. Puce à protéine, caractérisée en ce qu'elle contient au moins un polypeptide selon l'une des revendications 5 à 7, 32 à 47, 65 et 67, ou au moins un anticorps selon l'une des revendications 70 et 71, immobilisé sur le support de ladite puce.

77. Puce à protéine selon la revendication 76, caractérisée en ce qu'elle contient
25 en outre au moins un polypeptide de micro-organisme autre que *Streptococcus agalactiae* ou au moins un anticorps dirigé contre un composé de micro-organisme autre que *Streptococcus agalactiae*, immobilisé sur le support de ladite puce.

78. Kit ou nécessaire pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé,
30 caractérisé en ce qu'il comprend une puce à protéine selon l'une des revendications 76 et 77.

79. Kit ou nécessaire pour la détection et/ou l'identification d'un micro-organisme, caractérisé en ce qu'il comprend une puce à protéine selon la revendication 77.

80. Procédé de détection et/ou d'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé dans un échantillon biologique, caractérisé en ce qu'il met en œuvre une séquence nucléotidique selon l'une des revendications 2 à 4, 8, 9, 11 à 13, 17 à 25, 48 à 52 et 68.

5 81. Procédé selon la revendication 80, caractérisé en ce qu'il comporte les étapes suivantes :

- a) éventuellement, isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ;
- b) amplification spécifique de l'ADN de bactéries appartenant à l'espèce *Streptococcus*
10 *agalactiae* ou à un micro-organisme associé à l'aide d'au moins une amorce selon l'une des revendications 48 à 52 ;
- c) mise en évidence des produits d'amplification.

82. Procédé selon la revendication 80, caractérisé en ce qu'il comprend les étapes suivantes :

- 15 a) mise en contact d'une sonde nucléotidique selon l'une des revendications 48 à 52, avec un échantillon biologique, l'acide nucléique contenu dans l'échantillon biologique ayant, le cas échéant, préalablement été rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à l'acide nucléique d'une
20 bactérie appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé ;
- b) mise en évidence de l'hybride éventuellement formé entre la sonde nucléotidique et l'acide nucléique de l'échantillon biologique.

83. Procédé selon la revendication 80, caractérisé en ce qu'il comprend les étapes suivantes :

- 25 a) mise en contact d'une sonde nucléotidique immobilisée sur un support selon la revendication 50 avec un échantillon biologique, l'acide nucléique de l'échantillon ayant, le cas échéant, été préalablement rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à l'acide nucléique d'une bactérie appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé ;
- 30 b) mise en contact de l'hybride formé entre la sonde nucléotidique immobilisée sur un support et l'acide nucléique contenu dans l'échantillon biologique, le cas échéant après élimination de l'acide nucléique de l'échantillon biologique, n'ayant pas hybridé avec la sonde, avec une sonde nucléotidique marquée selon la revendication 49 ;

c) mise en évidence du nouvel hybride formé à l'étape b).

84. Procédé selon la revendication 83, caractérisé en ce que, préalablement à l'étape a), l'ADN de l'échantillon biologique ou l'ADNc obtenu éventuellement par transcription inverse de l'ARN de l'échantillon, est amplifié à l'aide d'au moins une
5 amorce selon l'une des revendications 48 à 52.

85. Kit ou nécessaire pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :

- a) une sonde nucléotidique selon l'une des revendications 48 à 52;
- 10 b) éventuellement, les réactifs nécessaires à la mise en œuvre d'une réaction d'hybridation ;
- c) éventuellement, au moins une amorce selon l'une des revendications 48 à 52 ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN.

86. Kit ou nécessaire pour la détection et/ou l'identification de bactéries
15 appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :

- a) une sonde nucléotidique, dite sonde de capture, selon la revendication 50 ;
- b) une sonde oligonucléotidique, dite sonde de révélation, selon la revendication 49;
- c) éventuellement, au moins une amorce selon l'une des revendications 48 à 52 ainsi
20 que les réactifs nécessaires à une réaction d'amplification de l'ADN.

87. Kit ou nécessaire pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae* ou à un micro-organisme associé, caractérisé en ce qu'il comprend les éléments suivants :

- a) au moins une amorce selon l'une des revendications 48 à 52;
- 25 b) éventuellement, les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN ;
- c) éventuellement, un composant permettant de vérifier la séquence du fragment amplifié, plus particulièrement une sonde oligonucléotidique selon l'une des revendications 48 à 52.

88. Procédé selon les revendications 72, 73 et 80 à 84 ou kit ou nécessaire selon les revendications 74, 78, 79 et 85 à 87 pour la détection et/ou l'identification de bactéries appartenant à l'espèce *Streptococcus agalactiae*, caractérisé en ce que ladite amorce et/ou ladite sonde sont choisies parmi les séquences nucléotidiques selon l'une des revendications 2 à 4, 8 à 28, 48 à 52, et 68 spécifiques de l'espèce *Streptococcus*

agalactiae, en ce que lesdits polypeptides sont choisis parmi les polypeptides selon l'une des revendications 5 à 7, 32 à 47, 65 et 67 spécifiques de l'espèce *Streptococcus agalactiae* et en ce que lesdits anticorps sont choisis parmi les anticorps selon l'une des revendications 70 et 71 dirigés contre les polypeptides choisis parmi les polypeptides selon l'une des revendications 5 à 7, 32 à 47, 65 et 67 spécifiques de l'espèce *Streptococcus agalactiae*.

89. Souche de *Streptococcus agalactiae*, caractérisée en ce qu'elle contient au moins une mutation dans au moins une séquence nucléotidique selon l'une des revendications 2 à 4, 8 à 28.

90. Souche de *Streptococcus agalactiae* selon la revendication 89, caractérisée en ce que la mutation mène à une inactivation du gène.

91. Souche de *Streptococcus agalactiae* selon la revendication 89, caractérisée en ce que la mutation mène à une surexpression du gène.

92. Utilisation d'une séquence nucléotidique selon l'une des revendications 2 à 4, 8 à 28, d'un polypeptide selon l'une des revendications 5 à 7, 32 à 47, 65 et 67 d'un anticorps selon l'une des revendications 70 et 71, d'une cellule selon l'une des revendications 60 à 62, et/ou d'un animal transformé selon la revendication 63 pour la sélection de composé organique ou inorganique capable de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, et/ou de modifier la réplication cellulaire de cellules eucaryotes ou procaryotes ou capables d'induire, d'inhiber ou d'aggraver chez un organisme animal ou humain les pathologies liées à une infection par *Streptococcus agalactiae* ou par un micro-organisme associé.

93. Méthode de sélection de composé capable de se lier à un polypeptide selon l'une des revendications 5 à 7, 32 à 47, 65 et 67, capable de se lier à une séquence nucléotidique selon l'une des revendications 2 à 4, 8 à 28, ou capable de reconnaître un anticorps selon l'une des revendications 70 et 71, et/ou capable de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, et/ou de modifier la réplication cellulaire de cellules eucaryotes ou procaryotes, ou capable d'induire, d'inhiber ou d'aggraver chez un organisme animal ou humain les pathologies liées à une infection par *Streptococcus agalactiae*, caractérisée en ce qu'elle comprend les étapes suivantes :

a) mise en contact dudit composé avec ledit polypeptide, ladite séquence nucléotidique, avec une cellule transformée selon l'une des revendications 60 à 62, et/ou administration dudit composé à un animal transformé selon la revendication 63 ;

- b) détermination de la capacité dudit composé à se lier avec ledit polypeptide ou ladite séquence nucléotidique, ou de moduler, de réguler, d'induire ou d'inhiber l'expression de gènes, ou de moduler la croissance ou la réplication cellulaire, ou d'induire, d'inhiber ou d'aggraver chez ledit organisme animal ou humain les pathologies liées à une infection par *Streptococcus agalactiae* ou par un micro-organisme associé.

94. Composition pharmaceutique comprenant un composé choisi parmi les composés suivants :

- a) une séquence nucléotidique selon l'une des revendications 2 à 4, 8 à 28 ;
b) un polypeptide selon l'une des revendications 5 à 7, 32 à 47, 65 et 67;
c) un vecteur selon la revendication 59 ou 69 ; et
d) un anticorps selon la revendication 70 ou 71.

95. Composition selon la revendication 94, éventuellement en association avec un véhicule pharmaceutiquement acceptable.

96. Composition pharmaceutique selon l'une des revendications 94 et 95 pour la prévention et le traitement d'une infection par une bactérie appartenant à l'espèce *Streptococcus agalactiae*.

97. Composition immunogène, caractérisée en ce qu'elle comprend un ou plusieurs polypeptides selon l'une des revendications 5 à 7, 32 à 47, 65, et/ou un ou plusieurs polypeptides hybrides selon la revendication 67.

98. Utilisation d'une cellule selon l'une des revendications 60 à 62, ou d'un vecteur selon l'une des revendications 59 ou 69 pour la préparation d'une composition vaccinale.

99. Composition vaccinale, caractérisée en ce qu'elle contient un polynucléotide selon l'une des revendications 1 à 4, 8 à 28, un vecteur selon l'une des revendications 59 ou 69, et/ou une cellule selon l'une des revendications 60 à 62.

100. Composition vaccinale, caractérisée en ce qu'elle contient au moins un polypeptide codé par un polynucléotide de séquence choisie parmi SEQ ID N°1503,678,2192,1861,584,280.

101. Composition vaccinale selon la revendication 100, caractérisée en ce qu'il s'agit d'une composition vétérinaire

102. Composition immunogène capable d'induire une réponse immunitaire cellulaire ou humorale pour la prévention ou le traitement d'une infection par une bactérie appartenant à l'espèce *Streptococcus agalactiae*, caractérisée en ce qu'elle

comprend une composition immunogène selon la revendication 97, ou une composition vaccinale selon la revendication 99 ou 100, en association avec un véhicule pharmaceutiquement acceptable et éventuellement un ou plusieurs adjuvants de l'immunité appropriés.

5 103. Banque génomique de *Streptococcus agalactiae* CIP 82.45 (ATCC 12403).

104. Banque d'ADN génomique selon la revendication 101, caractérisée en ce que ladite banque d'ADN est clonée dans un plasmide.

10 105. Banque selon la revendication 101 ou 102, caractérisée en ce qu'il s'agit de la banque déposée à la CNCM le 28 décembre 2000 sous le N° I-2610.

106. Utilisation des banques génomiques selon l'une des revendications 101 à 103 pour isoler des séquences nucléotidiques spécifiques de *Streptococcus agalactiae*, caractérisée en ce que les séquences nucléotidiques de *Streptococcus* autres que *Streptococcus agalactiae* CIP 82.45 (ATCC 12403) sont alignées et en ce que les
15 données obtenues par cet alignement sont traitées pour isoler lesdites séquences spécifiques.

107. Procédé d'identification de séquence spécifique de *Streptococcus agalactiae*, caractérisé par l'alignement de séquences nucléotidiques de *Streptococcus agalactiae* selon les revendications 1 à 4, 8 à 9 et le traitement des données obtenues par
20 cet alignement pour isoler les séquences spécifiques.

108. Souche mutante NEM 1979 de *Streptococcus agalactiae* selon la revendication 89 déposée à la CNCM le 24 avril 2002 sous le N° I-2861.

109. Souche mutante NEM 2056 de *Streptococcus agalactiae* selon la revendication 89 déposée à la CNCM le 24 avril 2002 sous le N° I-2862.

25 110. Souche mutante NEM 2057 de *Streptococcus agalactiae* selon la revendication 89 déposée à la CNCM le 24 avril 2002 sous le N° I-2863.

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
21 novembre 2002 (21.11.2002)

PCT

(10) Numéro de publication internationale
WO 02/092818 A3

(51) Classification internationale des brevets⁷ :

C12N 15/31, C07K

14/315, 16/12, C12N 15/63, A61K 39/09, 39/40, 31/711,
C12Q 1/68, G01N 33/53, A01K 671/027, A01H 5/00

(21) Numéro de la demande internationale :

PCT/IB02/03059

(22) Date de dépôt international : 26 avril 2002 (26.04.2002)

(25) Langue de dépôt :

français

(26) Langue de publication :

français

(30) Données relatives à la priorité :

2001/05642 26 avril 2001 (26.04.2001) FR

(71) Déposants (*pour tous les États désignés sauf US*) : **INSTITUT PASTEUR** [FR/FR]; 25-28, rue de la Glacière, F-75013 Paris (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel-Ange, F-75016 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : **GLASER, Philippe** [FR/FR]; 72 rue de la Glacière, F-75013 Paris (FR). **RUSNIOK, Christophe** [FR/FR]; 44 rue des Gallardons, F-92350 Le Plessis Robinson (FR). **CHEVALIER, Fabien** [FR/FR]; 29, avenue Léon Blum, F-94230 Cachan (FR). **FRANGEUL, Lionel** [FR/FR]; 2, rue Mouton Duvernet, F-75014 Paris (FR). **LALIOUI, Lila** [DZ/FR]; 7, rue des Presles, F-94170 Le Perreux sur Marne (FR). **ZOUINE, Mohamed** [FR/FR]; 4, avenue de Normandie, log. 156 91940, F- Les Ulis (FR). **COUVE, Elisabeth** [FR/FR]; 6, rue de Limours, F-78120 Rambouillet (FR). **BUCHRIESER, Carmen** [AT/FR]; 11, rue de l'Amiral Mouchez, F-75013 Paris (FR). **POYART, Claire** [FR/FR];

20, rue Léon Blum, F-92260 Fontenay-aux-Roses (FR). **TRIEU-CUOT, Patrick** [FR/FR]; 20, rue Léon Blum, F-92260 Fontenay aux Roses (FR). **KUNST, Frank** [FR/FR]; 46, rue Barbes, F-92200 Ivry sur Seine (FR).

(74) Mandataires : **MARTIN, Jean-Jacques** etc.; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) États désignés (*national*) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Déclaration en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement*

Publiée :

— *avec rapport de recherche internationale*
— *avec la partie réservée au listage des séquences de la description publiée séparément sous forme électronique et disponible sur demande auprès du Bureau international*

(88) Date de publication du rapport de recherche
internationale: 28 août 2003

[Suite sur la page suivante]

(54) Title: STREPTOCOCCUS AGALACTIAE GENOME SEQUENCE, USE FOR DEVELOPING VACCINES, DIAGNOSTIC TOOLS, AND FOR IDENTIFYING THERAPEUTIC TARGETS

(54) Titre : SEQUENCE DU GENOME STREPTOCOCCUS AGALACTIAE ET SES UTILISATIONS

(57) Abstract: The invention concerns the genome sequence and nucleotide sequences coding for *Streptococcus agalactiae* polypeptides, such as cellular envelope polypeptides, or secreted or specific polypeptides, or polypeptides involved in the metabolism and the replication process, as well as vectors or cells comprising said sequences. The invention also concerns the use thereof for developing vaccines, diagnostic tools, DNA chips and for identifying therapeutic targets.

(57) Abrégé : L'invention a pour objet la séquence génomique et des séquences nucléotidiques codant pour des polypeptides de *Streptococcus agalactiae*, tels que des polypeptides d'enveloppe cellulaire, ou des polypeptides sécrétés ou spécifiques, ou impliqués dans le métabolisme et dans le processus de réplication, ainsi que des vecteurs ou cellules incluant lesdites séquences. L'invention porte également sur leur application au développement de vaccins, d'outils de diagnostic, de puces à ADN et à l'identification de cibles thérapeutiques.



WO 02/092818 A3



En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 02/03059

A. CLASSIFICATION OF SUBJECT MATTER		
C12N15/31 C07K14/315 C07K16/12 C12N15/63 IPC 7 A61K39/09 A61K39/40 A61K31/711 C12Q1/68 G01N33/53 A01K67/027 According to International Patent Classification (IPC) or to both national classification and IPC A01H5/00		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, MEDLINE, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>SPELLERBERG B ET AL: "LMB, A PROTEIN WITH SIMILARITIES TO THE LRAI ADHESIN FAMILY, MEDIATES ATTACHMENT OF STREPTOCOCCUS AGALACTIAE TO HUMAN LAMININ" INFECTION AND IMMUNITY, AMERICAN SOCIETY FOR MICROBIOLOGY. WASHINGTON, US, vol. 67, no. 2, February 1999 (11.02.99), pages 871-878, XP000973065 ISSN: 0019-9567 -& DATABASE EMBL [en ligne] 11 February 1999 (11.02.99) SPELLERBERG B. ET AL.: "Streptococcus agalactiae Lmb (lmb) gene, complete cds; and unknown gene" Database accession no. AF062533 XP002221154 The whole document</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	<p>1-9,12, 31, 48-99, 102, 107-110</p>
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>		
Date of the actual completion of the international search 18 November 2002 (18.11.02)		Date of mailing of the international search report 19 February 2003 (19.02.03)
Name and mailing address of the ISA/ S.P.T.O.		Authorized officer
Facsimile No.		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 02/03059

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 01 14421 A (MEDIMMUNE, INC.) 1 March 2001 (01.03.01)</p> <p>page 11, line 13 -page 26, line 20 Sequence listing SEQ ID NO:5, 6</p>	1-9,12, 31, 48-99, 102, 107-110
X	<p>--- DATABASE SWALL [in line] 1 March 2001 (01.03.01) DE BOEVER, E.H. ET AL.: "TraG-related protein" Database accession no. Q9F1G0 XP002221155 The whole document & ERIKA H. DE BOEVER ET AL.: "Enterococcus faecalis conjugative plasmid pAM373: complete nucleotide sequence and genetic analysis of sex phromone response" MOLECULAR MICROBIOLOGY, vol. 37, no. 6, 2000, pages 1327-1341,</p>	5,7,8
A	<p>--- DORAN, T. I. ET AL.: "Factors Influencing Release of Type III Antigens by Group B Streptococci" INFECT. IMMUN., February 1981 (02,02,81) vol. 31, no. 2, pages 615-623, XP002191322 page 621, hand left column , alinéa 2 - alinéa 3 -----</p>	1-102, 107-110

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 02/03059

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see supplementary sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

In particular 1-102, 107-110

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

The International Searching Authority has determined that the present international application contains multiple (groups of) inventions, namely:

1. Claims : (in part) 1-102, 107-110

Nucleotide sequence of *Streptococcus agalactiae*, SEQ ID No. 1, fragments and homologues; polypeptides coded by said sequences, derived polypeptides, hybrids and antibodies; probe and primer derived from the nucleotide sequence; use thereof for selecting compounds having an effect on disease states caused by an *S. agalactiae* infection; pharmaceutical compositions; DNA chips and protein; kit containing said chips; cloning vector, host cell, plant or animal containing said nucleotide sequence; use of the antibodies and of the nucleotide sequences to identify *Streptococcus agalactiae*; strain of *S. agalactiae* containing a mutation in the sequence SEQ ID No. 1; use of the nucleotide sequences to identify *S. agalactiae*-specific sequences.

2. Claims: (in part) 1-102, 107-110

Same as invention n° 1, for the nucleotide sequences SEQ ID Nos. 2-139, 2345 and 4482-6617 respectively.

3. Claims: 103-106

Gene library of *Streptococcus agalactiae* and use thereof.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 02/03059

Patent document cited in search report		Publication date	Patent familiy member(s)		Publication date
WO 0114421	A	01-03-2001	AU	7076100 A	19-03-2001
			EP	1210366 A1	05-06-2002
			WO	0114421 A1	01-03-2001

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/IB 02/03059

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 7	C12N15/31	C07K14/315	C07K16/12	C12N15/63	A61K39/09
	A61K39/40	A61K31/711	C12Q1/68	G01N33/53	A01K67/027
	A01H5/00				

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 C12N C07K A61K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, MEDLINE, EMBL

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>SPELLERBERG B ET AL: "LMB, A PROTEIN WITH SIMILARITIES TO THE LRAI ADHESIN FAMILY, MEDIATES ATTACHMENT OF STREPTOCOCCUS AGALACTIAE TO HUMAN LAMININ" INFECTION AND IMMUNITY, AMERICAN SOCIETY FOR MICROBIOLOGY. WASHINGTON, US, vol. 67, no. 2, février 1999 (1999-02), pages 871-878, XP000973065 ISSN: 0019-9567 -& DATABASE EMBL [en ligne] 11 février 1999 (1999-02-11) SPELLERBERG B. ET AL.: "Streptococcus agalactiae Lmb (lmb) gene, complete cds; and unknown gene" Database accession no. AF062533 XP002221154 le document en entier ---</p>	<p>1-9,12, 31, 48-99, 102, 107-110</p>
	-/-	

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

18 novembre 2002

Date d'expédition du présent rapport de recherche internationale

19. 02. 2003

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

MONTERO LOPEZ B.

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/IB 02/03059

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>WO 01 14421 A (MEDIMMUNE, INC.) 1 mars 2001 (2001-03-01)</p> <p>page 11, ligne 13 -page 26, ligne 20 Sequence listing SEQ ID NO:5, 6</p> <p>---</p>	<p>1-9,12, 31, 48-99, 102, 107-110</p>
X	<p>DATABASE SWALL [en ligne] 1 mars 2001 (2001-03-01) DE BOEVER, E.H. ET AL.: "TraG-related protein" Database accession no. Q9F1G0 XP002221155 le document en entier & ERIKA H. DE BOEVER ET AL.: "Enterococcus faecalis conjugative plasmid pAM373: complete nucleotide sequence and genetic analysis of sex phromone response" MOLECULAR MICROBIOLOGY, vol. 37, no. 6, 2000, pages 1327-1341,</p> <p>---</p>	<p>5,7,8</p>
A	<p>DORAN, T. I. ET AL.: "Factors Influencing Release of Type III Antigens by Group B Streptococci" INFECT. IMMUN., vol. 31, no. 2, février 1981 (1981-02), pages 615-623, XP002191322 page 621, colonne de gauche, alinéa 2 - alinéa 3</p> <p>-----</p>	<p>1-102, 107-110</p>

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/IB. 02/03059

Cadre I Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n^{os} se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:
2. ☐ Les revendications n^{os} se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:
3. ☐ Les revendications n^{os} sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.
2. ☐ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.
3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n^{os}
4. ☒ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n^{os}
Partiellement 1-102, 107-110

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: Partiellement 1-102, 107-110

Séquence nucléotidique de *Streptococcus agalactiae* SEQ ID NO:1, fragments et homologues; polypeptides codés par ces séquences, polypeptides dérivés, hybrides et anticorps; sonde et amorce dérivés de la séquence nucléotidique; utilisation de ceux-ci pour la sélection de composés ayant une influence sur les pathologies liées à une infection par *S. agalactiae*; compositions pharmaceutiques; puces à ADN et protéine; kit ou nécessaire contenant lesdites puces; vecteur de clonage, cellule hôte, végétal ou animal contenant la séquence nucléotidique; utilisation des anticorps et des séquences nucléotidiques pour l'identification de *Streptococcus agalactiae*; souche de *S. agalactiae* contenant une mutation dans la séquence SEQ ID NO:1; utilisation des séquences nucléotidiques pour identifier des séquences spécifiques de *S. agalactiae*.

2. revendications: Partiellement 1-102, 107-110

Idem au sujet 1 pour, respectivement les séquences nucléotidiques SEQ ID NOs:2-139, 2345, and 4482-6617

3. revendications: 103-106

Banque génomique de *Straptococcus agalactiae* et son utilisation

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/IB 02/03059

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 0114421 A	01-03-2001	AU 7076100 A	19-03-2001
		EP 1210366 A1	05-06-2002
		WO 0114421 A1	01-03-2001

VERSION CORRIGÉE

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
21 novembre 2002 (21.11.2002)

PCT

(10) Numéro de publication internationale
WO 2002/092818 A3

(51) Classification internationale des brevets⁷ :

C12N 15/31, C07K

14/315, 16/12, C12N 15/63, A61K 39/09, 39/40, 31/711,
C12Q 1/68, G01N 33/53, A01K 671/027, A01H 5/00

(21) Numéro de la demande internationale :

PCT/IB2002/003059

(22) Date de dépôt international : 26 avril 2002 (26.04.2002)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :

2001/05642 26 avril 2001 (26.04.2001) FR

(71) Déposants (pour tous les États désignés sauf US) : **INSTITUT PASTEUR** [FR/FR]; 25-28, rue du Docteur Roux, F-75015 Paris (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel-Ange, F-75016 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **GLASER, Philippe** [FR/FR]; 72 rue de la Glacière, F-75013 Paris (FR). **RUSNIOK, Christophe** [FR/FR]; 44 rue des Gallardons, F-92350 Le Plessis Robinson (FR). **CHEVALIER, Fabien** [FR/FR]; 29, avenue Léon Blum, F-94230 Cachan (FR). **FRANGEUL, Lionel** [FR/FR]; 2, rue Mouton Duvernet, F-75014 Paris (FR). **LALIOUI, Lila** [DZ/FR]; 7, rue des Presles, F-94170 Le Perreux sur Marne (FR). **ZOUINE, Mohamed** [FR/FR]; 4, avenue de Normandie, log. 156 91940, F- Les Ulis (FR). **COUVE, Elisabeth** [FR/FR]; 6, rue de Limours, F-78120 Rambouillet (FR). **BUCHRIESER, Carmen** [AT/FR]; 11, rue de l'Amiral Mouchez, F-75013 Paris (FR). **POYART, Claire** [FR/FR];

20, rue Léon Blum, F-92260 Fontenay-aux-Roses (FR). **TRIEU-CUOT, Patrick** [FR/FR]; 20, rue Léon Blum, F-92260 Fontenay aux Roses (FR). **KUNST, Frank** [FR/FR]; 46, rue Barbes, F-92200 Ivry sur Seine (FR).

(74) Mandataires : **MARTIN, Jean-Jacques** etc.; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) États désignés (régional) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Déclaration en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement

Publiée :

— avec rapport de recherche internationale
— avec la partie réservée au listage des séquences de la description publiée séparément sous forme électronique et disponible sur demande auprès du Bureau international

(88) Date de publication du rapport de recherche internationale:

28 août 2003

[Suite sur la page suivante]

(54) Title: STREPTOCOCCUS AGALACTIAE GENOME SEQUENCE, USE FOR DEVELOPING VACCINES, DIAGNOSTIC TOOLS, AND FOR IDENTIFYING THERAPEUTIC TARGETS

(54) Titre : SEQUENCE DU GENOME STREPTOCOCCUS AGALACTIAE ET SES UTILISATIONS

(57) Abstract: The invention concerns the genome sequence and nucleotide sequences coding for *Streptococcus agalactiae* polypeptides, such as cellular envelope polypeptides, or secreted or specific polypeptides, or polypeptides involved in the metabolism and the replication process, as well as vectors or cells comprising said sequences. The invention also concerns the use thereof for developing vaccines, diagnostic tools, DNA chips and for identifying therapeutic targets.

(57) Abrégé : L'invention a pour objet la séquence génomique et des séquences nucléotidiques codant pour des polypeptides de *Streptococcus agalactiae*, tels que des polypeptides d'enveloppe cellulaire, ou des polypeptides sécrétés ou spécifiques, ou impliqués dans le métabolisme et dans le processus de réplication, ainsi que des vecteurs ou cellules incluant lesdites séquences. L'invention porte également sur leur application au développement de vaccins, d'outils de diagnostic, de puces à ADN et à l'identification de cibles thérapeutiques.



WO 2002/092818 A3



(48) Date de publication de la présente version corrigée:
4 mars 2004

(15) Renseignements relatifs à la correction:
voir la Gazette du PCT n° 10/2004 du 4 mars 2004, Section II

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.